SCALING PATTERNS AND DRIVERS OF SPECIES RICHNESS AND TURNOVER ACROSS THE AFROTROPICS

CURTIS JUNIOR BAILEY
(BSc HONOURS GEOINFORMATICS)

Thesis presented in fulfilment of the requirements for the degree of Master of Science in the Faculty of Science at Stellenbosch University

Supervisor: Dr Helen De Klerk
Co-Supervisors: Prof. Cang Hui & Mrs Zahn Munch

March 2017

DECLARATION

By submitting this report electronically, I declare that the entirety of the work contained herein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously, in its entirety or in part submitted it for obtaining any qualification.

Date: March 2017

ABSTRACT

Understanding the broad scale patterns and environmental determinants of amphibian richness (α-diversity) and turnover (β-diversity) is becoming ever more crucial as communities, ecosystems and landscapes are increasingly threatened by anthropogenic processes and climate change. Spatial scale, the grain at which sampling takes place or analysis occurs, confounds understanding of such diversity patterns and the relationships these exhibit with environmental processes. This is because various processes operate at different spatial scales resulting in different patterns emerging at different spatial scales. This is problematic, because patterns, and pattern-process' relationships which emerge at one scale may not hold at another scale thereby confounding our understanding of how biotic patterns are generated and maintained, thus leading to misguided conservation strategies and policies. The focus of the present study was thus to examine the relationship between present patterns of anuran richness and turnover, the relationship of these with several contemporary environmental processes, and how these patterns and pattern-process relationships are influenced by spatial grain.

Using IUCN range data for the Afrotropical region, amphibian richness and turnover patterns were generated using a Geographic Information System and quantified using the recently formulated zeta diversity partitioning method (ζ). These patterns were then related to several contemporary environmental variables/processes hypothesised to govern amphibian diversity across a range of spatial scales using both the global Ordinary Least Squares (OLS) method and local spatial Geographically Weighted Regression (GWR) across three nested spatial grains.

The results show that amphibian richness and turnover patterns (visual & empirical) and the relationships these exhibit with environmental conditions were sensitive to spatial grain. Visually, spatial patterns were more distinct at finer grains, but visibly smooth at the coarser grains. The statistics describing these patterns suggest that the scaling behaviour follows an increasing linear and nonlinear trend across spatial grain. Results from both the OLS and GWR models confirm that contemporary environmental conditions are significant determinants of both anuran richness and turnover patterns across spatial grain. Precipitation was the strongest determinant of anuran richness while topographic complexity best explained turnover. The relationship between both diversity components and environmental conditions, however, was scale dependent with environmental conditions explaining a greater proportion of the variation in these biotic patterns at coarser than finer grains. Studies conducted across multiple grains is thus recommended to improve current understanding of biodiversity patterns and the relationship they exhibit with environmental processes.

Keywords: Amphibians, α-diversity, β-diversity, environment, spatial scale, spatial grain, Geographic Information System, regression

OPSOMMING

Dit word al hoe meer noodsaaklik om die grootskaalse patrone en omgewingsdeterminante van amfibiese rykdom (α-diversiteit) en omset (β-diversiteit) te verstaan soos gemeenskappe, ekosisteme en landskappe bedreig word deur mensgemaakte prosesse en klimaatsverandering. Ruimetelike skaal, die grein waarteen monstering of analise plaasvind, verydel die begrip van sulke diversiteitspatrone en die verwantskappe wat vertoon word met omgwewingsprosesse. Dit is omdat verskillende prosesse teen verskillende ruimtelike skale funksioneer wat daartoe lei dat verkillende patrone na vore kom teen verskillende ruimtelike skale. Dit is problematies omdat patrone en patroonproses verwantskappe wat teen ‘n sekere skaal na vore kom nie noodwendig teen ‘n ander skaal geld nie. Dit verydel die begrip van hoe biotiese gegenereer en in stand gehou word. En lei daartoe dat misleidende bewaringsstrategieë en -beleide geïmplementeer word. Die fokus van hierdie studie was dus om die verwantskap tussen huidige patrone van amfibiese rykheid en omset, die verhouding hiervan met verskeie kontemporêre omgewingsprosesse, en hoe hierdie patrone en patroonproses verhoudings in stand gehou word oor die ruimtelike grein te ondersoek.

IUCN data vir die Afrotropiese gebied is gebruik om amfibiese rykheid en omsetpatrone te ondersoek met behulp van Geografiese Inligtingstelsels (GIS) en die onlangs-geformuleerde zeta diversiteit partisioneringstegniek (ζ). Hierdie patrone is dan verbind met verskeie kontemporêre omgewingsveranderlikes of omgwingsprosesse, wat veronderstel is om amfibiese diversiteit te reguleer, oor verskeie ruimtelike skale deur gebruik te maak van gewone kleinste kwadrate (GKK) en plaaslike geografies-geweegde regressie (GGR) oor drie ruimtelile skale.

Die resultate toon dat amfibiese rykheid en omsetpatrone (visueel en empiries), en die verwantskap wat dit toon tot omgewingstoestande, sensitief is vir ruimtelike grein. Ruimtelike patrone was meer opmerklik teen ‘n fyner grein, maar word meer geleideli/gladder teen ‘n growwer grein. Die statistiek van die patrone dui daarop dat die skaleringseienskappe ‘n verhoogde lineêre-, mags- en eksponensiële neiging volg met verhoogde ruimtelike grein resultate van beide die GKK en GGR modelle bevestig dat kontemporêre omgewingstoestande beduidende determinante is van amfibiese rykheid en omsetpatrone. Reënval was die sterkste determinant van amfibiese rykheid en topografiese kompleksiteit het die beste verklaring gebied vir omset. Die verwantskap tussen diversiteitskomponente en omgewingstoestande was afhanklik van skaal met omgewingstoestande wat ‘n groter gedeelte van die afwyking in die biotiese patrone teen growwer as teen fyner greine verklaar. Dit word dus aanbeveel dat

toekomstige studies wat diversiteit patrone en hul verwantskap met omgewingsprosesse ondersoek so doen oor verskeie ruimtelik skale om die huidiglike begrip daarvan te verbeter.

Sleutelwoorde: Amfibieë, α-diversiteit, β-diversiteit, omgewing, ruimtelike skaal, ruimtelike grein, Geografiese inligingstelsels, regressie.

ACKNOWLEDGEMENTS

I sincerely thank:

- **Dr Helen De Klerk**, my main supervisor for her expertise, guidance and devotion to the project. You were an exceptional mentor, always pushing me to excel and do my best;
- **Mrs Zahn Munch**, my co-supervisor for her guidance, technical assistance, motivational talks and perennial support. I was very blessed to have you as part of the project;
- **Prof Cang Hui** for his guidance, analytical eye and financial support;
- **Mrs Selene Ortell** for her constant support, the motivational talks, hugs and laughs;
- My **family**, for their undying love, the phone calls and encouraging messages;
- My **mom** who motivated me through the hard times and tears;
- My **friends** and **fellow Masters** students for the laughs and talks in the lab;
- **Marlin** for his incredible support and consistent best wishes over the past two years;
- The staff at Stellenbosch University's **Department of Geography and Environmental Studies**, in particular **Mrs Catherine Liederman, Prof Ronny Donaldson, Prof Adriaan Van Niekerk and Prof Sanette Ferreirra**. Your encouragement was salient through the hard times and much appreciated;
- Those who I did not mention but contributed to my experience as a student at Stellenbosch University;
- The **National Research Foundation** (NRF), **Centre for Invasive Biology** (CIB), and **Stellenbosch University (SU)** for the financial assistance and support. Without your generation contributions I would not have been able to complete my postgraduate studies;
- Lastly, **God** my creator and provider to whom I am eternally grateful for the endless opportunity I have been blessed with.

CONTENTS

TABLES

FIGURES

EQUATIONS

APPENDICES

ACRONYMS AND ABBREVIATIONS

α – diversity	Alpha diversity
β – diversity	Beta diversity
γ – diversity	Gamma diversity
AET	Actual evapotranspiration
AR	Autoregressive Model
CCA	Canonical Correspondence Analysis
CN	Condition Number
CAR	Conditional Autoregressive Model
DCA	Detrended Correspondence Analysis
GAM	General Additive Model
GAAU	Global Amphibian Assessment Unit
GDM	Generalised Dissimilarity modelling
GIS	Geographic Information System
GWR	Geographically Weighted Regression
EOC	Extent Of Occurrence
IUCN	International Union for the Conservation of Nature
OLS	Ordinary Least Squares
SAR	Simultaneous Autoregressive Model
VIF	Variance Inflation Factor
MIH	More Individual Hypothesis
MLR	Multiple Linear Regression
Mts.	Mountains
NDVI	Normalized Difference Vegetation Index
NPP	Net primary productivity
SRTM	Shuttle Radar Topography Mission
SAC	Spatial autocorrelation
WWF	World Wildlife Fund

CHAPTER 1: SETTING THE SCENE

"…the problem of relating phenomena across scales is the central problem of all of biology, in fact, all of science" – ***Levin (1992)***

1.1 INTRODUCTION

Species diversity patterns are the result of biotic and abiotic processes acting independently and interactively operating across a range of spatial scales. A primary objective of biogeography is to disentangle these processes to establish a link between pattern and process, so that the true spatial scales at which various relationships operate, may be observed (Fortin & Dale 2005; Hui 2015).

Environmental processes, in particular, are though - to exert profound influences on present distributions of biotic diversity including inventory diversity species composition at a locality (α-diversity) or across a region (γ-diversity) as well as the change in species composition between localities, namely, beta diversity (β-diversity). Several competing species-environment hypotheses including climate, climatic stability, energy availability, ecosystem productivity and physiographic complexity, have been proposed to explain broad-scale diversity patterns but the mechanisms underpinning these relationships remain unclear (Mcdonald et al. 2005; Clarke & Gaston 2006; Jetz et al. 2009; Diniz-filho et al. 2010; Saito et al. 2015).

In the past, much emphasis has been placed on patterns of inventory diversity and the associated environmental and evolutionary mechanisms underlying their origin and maintenance while diversity's differential component, beta diversity, has largely been negated. Recently, the paradigm has shifted towards beta diversity because it is central to an array of ecological topics including the mechanisms which form biotas, the delineation of biotic transition zones and selection of conservation sites (Jurasinski et al. 2009, 2012; Melo, Rangel & Diniz-Filho 2009). Beta diversity quantifies the change in species composition across geographic space and was initially measured as the ratio of regional to local diversity ($\beta = \gamma/\alpha$). Progressively, however, the convention has become to use compositional indices of dissimilarity, such as those of Jaccard, Sørensen and recently Zeta diversity partitioning, to quantify beta diversity (Diserud & Ødegaard 2007; Di Virgilio et al. 2014; Ricotta & Pavoine 2015).

Some processes and mechanisms underlying the distribution of alpha diversity are thought to drive gradients of beta diversity, though the degree and extent remain unclear. While several regional and broad-scale analysis have been conducted, results rendered have been largely conflicting (Arita & Rodriguez 2002; Buckley & Jetz 2007, 2008; Tuomisto 2010a; Barton et al.

2013) and the elucidating patterns and determinants of beta diversity remain largely unexplored and inadequately understood (Chen et al. 2011; Barwell et al. 2015).

A pervasive confounding factor for understanding diversity patterns and the relationships these exhibit with environmental conditions is spatial scale. Spatial scale may refer to any one, or combination, of several concepts (Ewers & Didham 2008). In macroecology, grain and extent are considered, where grain refers to spatial resolution, or size of the sampling unit used to measure diversity (Francis & Klopatek 2000) and extent, to the outer most bounds of a dataset (i.e. the size of landscape). Species diversity pattersn exhibit different patterns at different spatial scales, indicating a scale multiplicity relationship. Scale is thus crucial in understanding diversity pattern-processes interactions.

Much research has been done to try and establish how scale influences not only the visual perception of diversity patterns but also the metrics associated with describing these. Currently, most regional species diversity studies are conducted using a single scale of analysis (). This is problematic, since ecological pattern-processes occurring at one scale may not be representative of or applicable to those occurring at an another scale (Wu 2004). Ecologists are thus scaling up, not only in order to understand the complex and dynamic interactions between land and atmosphere, but also to try and convey information about fine scale ecological patterns and processes to broad scale applications (Francis & Klopatek 2000).

Cross scaling studies may part curly be useful to Amphibians, a taxa of prime conservation concern owing to their extreme sensitivity to environmental conditions. Recent precipitous declines in amphibian populations and localised extinctions have sparked global conservation initiatives prioritising the identification of climatic and anthropogenic threats to reduce or eradicate species losses (Werner et al. 2007; Sodhi et al. 2008). The urgency to link amphibian diversity patterns to those environmental processes hypothesised to govern their distribution has thus become ever more salient.

Proper ecosystem management for amphibians, however, can only be implemented if the processes and underlying mechanisms which create, maintain and confine their distributions are adequately understood. Therefore, as anthropogenic activity continues to alter and put strain on the natural environment, the need to evaluate the relationship between biotic patterns and the abiotic processes generating them across a spatial scale becomes ever more salient. Insights gained from such evaluations may lead to more rigid and effective applied conservation strategies, which may prove paramount to the survival of anurans and other taxa in the wake of impending climate and environmental change.

1.2 PROBLEM STATEMENT

Amphibian species and populations are declining at alarming rates worldwide (Kiesecker, Blaustein & Belden 2001; Sodhi et al. 2008). To reduce and eradicate species losses, effective and robust conservation strategies are needed. Such conservation efforts require an understanding of the mechanisms which shape and govern the distribution of biotic diversity patterns including: alpha diversity (α-diversity) - the number of species at locality and beta diversity (β-diversity, turnover) - the change in species composition between localities (Legendre 2008). Regional climate and contemporary environmental processes have been identified as amongst the most important determinants and drivers of broad scale biotic patterns (Gaston 2000). Accordingly, numerous studies have been conducted which have examined the relationship between biotic diversity and the environment for different taxa, across different parts of the world, using an array of statistical procedures have been conducted (Veech & Crist 2007; Hof et al. 2011; Dobrovolski et al. 2012).

A pervasive confounding factor for understanding the relationship between species diversity and environmental processes, however, is spatial scale (Lennon et al. 2001). Scale may refer to any one, or combination, of several concepts. In landscape ecology grain and extent are considered the most important where grain refers to the size of the sampling unit used to measure diversity, and extent to the outer most bounds of the study area (i.e. the size of the sampling extent). Scaling, a separate concept to scale, refers to the empirical transfer of information from one scale to another (Barton et al. 2013). Different spatial patterns emerge at different spatial scales owing to different processes (i.e. pattern-process interactions) operating at different spatial scales (Wu 2004).

Currently, most diversity pattern-environmental process relationships studies are conducted using only a single scale of analysis (Francis & Klopatek 2000; Hui 2009; Calderón-Patrón et al. 2013). This is problematic because biotic patterns are manufactured and maintained by multiple processes operating across range of spatial scales (Arita & Rodrıguez 2002). Results obtained and inferences made from single scale studies may thus not be completely reliable or accurate, but be artefacts of the scale of analysis.

In order to disentangle how patterns-processes interact, operate, and shape the natural environment, multiple scale or scaling studies are required. Although the need for multi-scale has been well established in the literature (Gibson & Ostrom 1998; Burgess et al. 2002; Legendre 2008; Hui & McGeoch 2014) analysis conducted at a single scale has remained the norm. Studies conducted across multiple scales may reveal how patterns and pattern-processes are affected by spatial scale.

Fortunately, the rapid technological advancements in Geographical Information Systems (GIS), which has been driven by need for the development of conceptual and predictive models for understanding Earths system processes, has come to the aid of ecologists (Wu 2004). GIS now offer an array of techniques and tools that are useful in the study of ecological patterns and processes across space, time and scale (Marceau DJ & Hay GJ 1999).

The intention of this thesis is thus to examine patterns of amphibian diversity, including alpha and beta diversity (turnover), how current environmental conditions have differentially influenced these and how these patterns and pattern-process relationships are influenced by spatial scale, specifically spatial grain.

1.3 RESEARCH AIM AND OBJECTIVES

The overarching aim of this study was to examine the relationship between the patterns and environmental drivers of anuran alpha and beta diversity (i.e. turnover) and how these are affected by spatial scale.

The aim was met by setting the following objectives:

1. To review literature pertaining to concepts salient to the study.
2. To devise an effective, robust and flexible research methodology.
3. To identify, collect and source the relevant data.
4. To explore and manipulate the data.
5. To map spatial patterns of alpha and beta diversity.
6. To explore the environmental determinants of alpha and beta diversity (i.e. pattern-process interactions).
7. To explore the relationship between environmental determinants of alpha and beta diversity (i.e. pattern-process interactions) at various spatial scales.

1.4 RESEARCH METHODOLOGY

The present study followed a quantitative research methodology. According to Getis (1999), the quantitative research approach using a GIS comprises six fundamental elements: data collection, data manipulation, data exploration, model specification, modelling and reporting (Figure 1.1).

In the data collection phase the relevant data are identified, collected and sourced. The data are then pre-processed and prepared for statistical analysis in the data manipulation phase. In the data exploration phase, the nature of the data should be inspected. This also comprises the final phase in which the data can be prepared for modelling.

In the model specification phase, variables are statistically evaluated with salient variables retained for modelling. In the modelling phase, hypotheses are tested using parametric or non-parametric tools and techniques. Proposed hypotheses are accepted or rejected in the reporting writing phase based on the results obtained from the modelling phase. Here, interpretation is also made with respect to existing literature. Getis (1999) notes that the

quantitative GIS method deviates from the traditional quantitative method because results are communicated spatially (i.e. through maps).

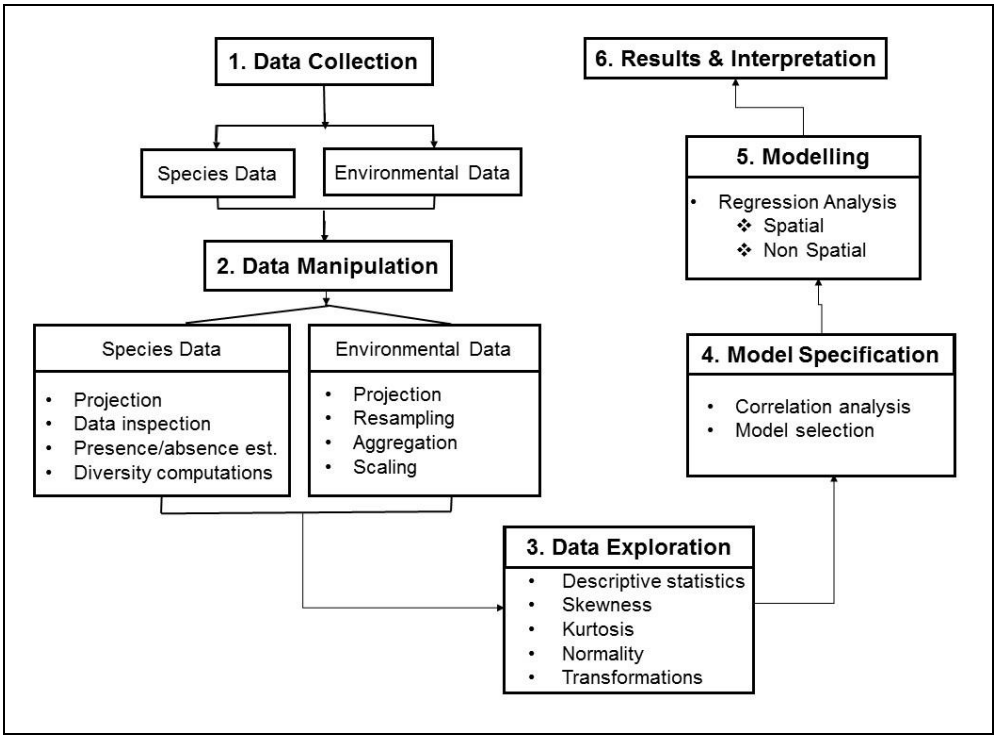


Figure 1.1: Research methodology adopted for empirical analysis in the study.

1.5 RESEARCH DESIGN

The research design (Figure 1.2) presents all the steps followed in the study in chronological order to effectively address the research problem. The research was executed in five phases presented as chapters. In phase one (Chapter 1), a research problem was identified which lead to the conceptualisation and development of the research aims and objects. In phase two (Chapter 2), a literature review was conducted to gain conceptual understanding of concepts, terms and methods central to the study. In phase three (Chapter 3), methods for the study were devised, data were collected and analyses were conducted to meet the objectives set for analytical component one; to map the patterns of alpha and beta diversity and relate these to current environmental correlates. In the penultimate phase (Chapter 4), the methods and analyses devised for analytical component one were repeated, but across multiple spatial scales. In phase 5 (Chapter 5), the main findings of the research, as observed in the two analytical components, were synthesised and summarised.

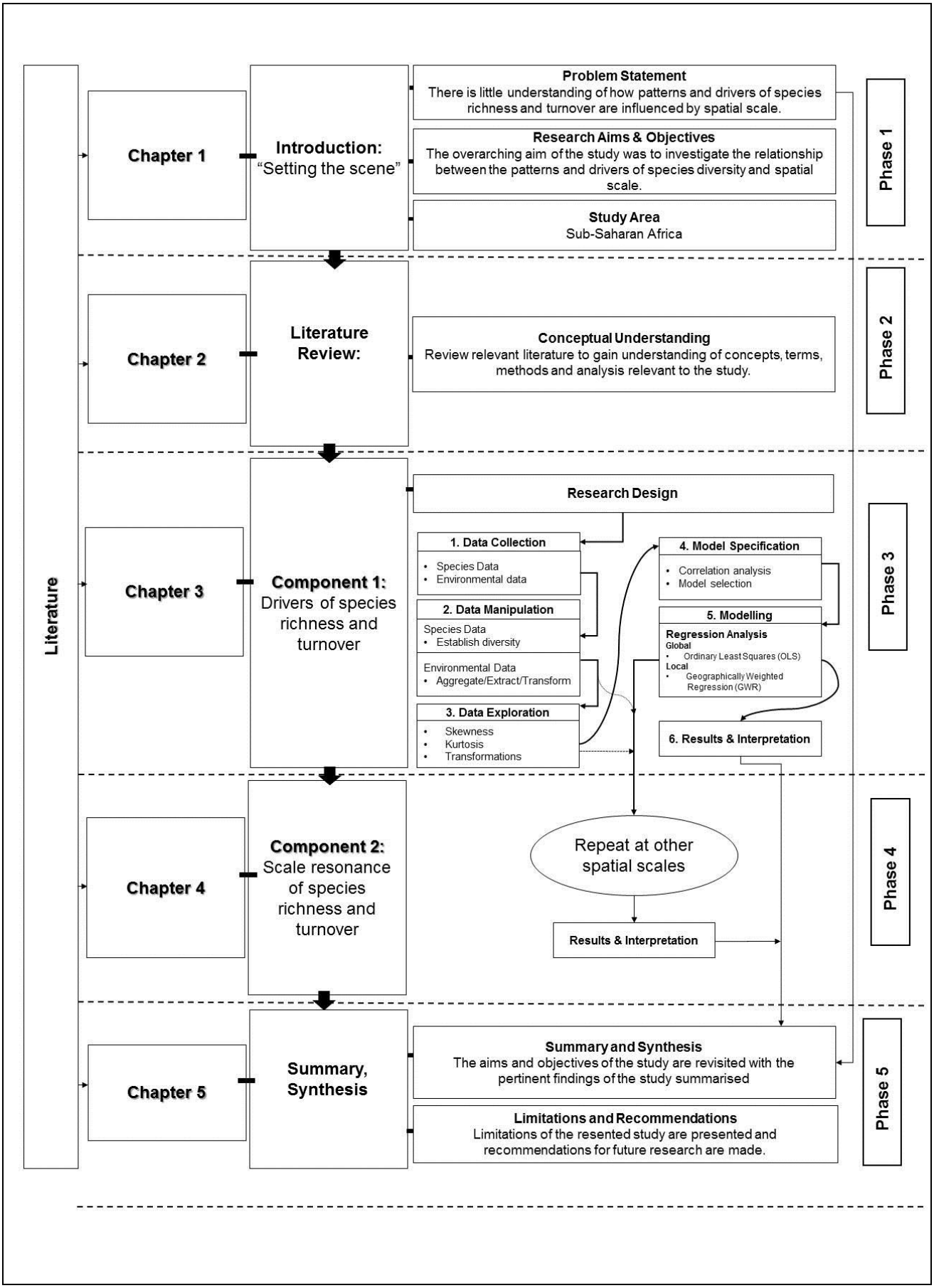


Figure 1.2: Research design followed for investigating the patterns and drivers of alpha and beta diversity across multiple spatial scales.

1.6 STUDY AREA

The study area (Figure 1.3) comprises continental Africa south of 15° north latitude (De Klerk et al. 2002; Linder et al. 2012; Qian & Ricklefs 2012) referred to as the Afrotropical realm or the Afrotropics.

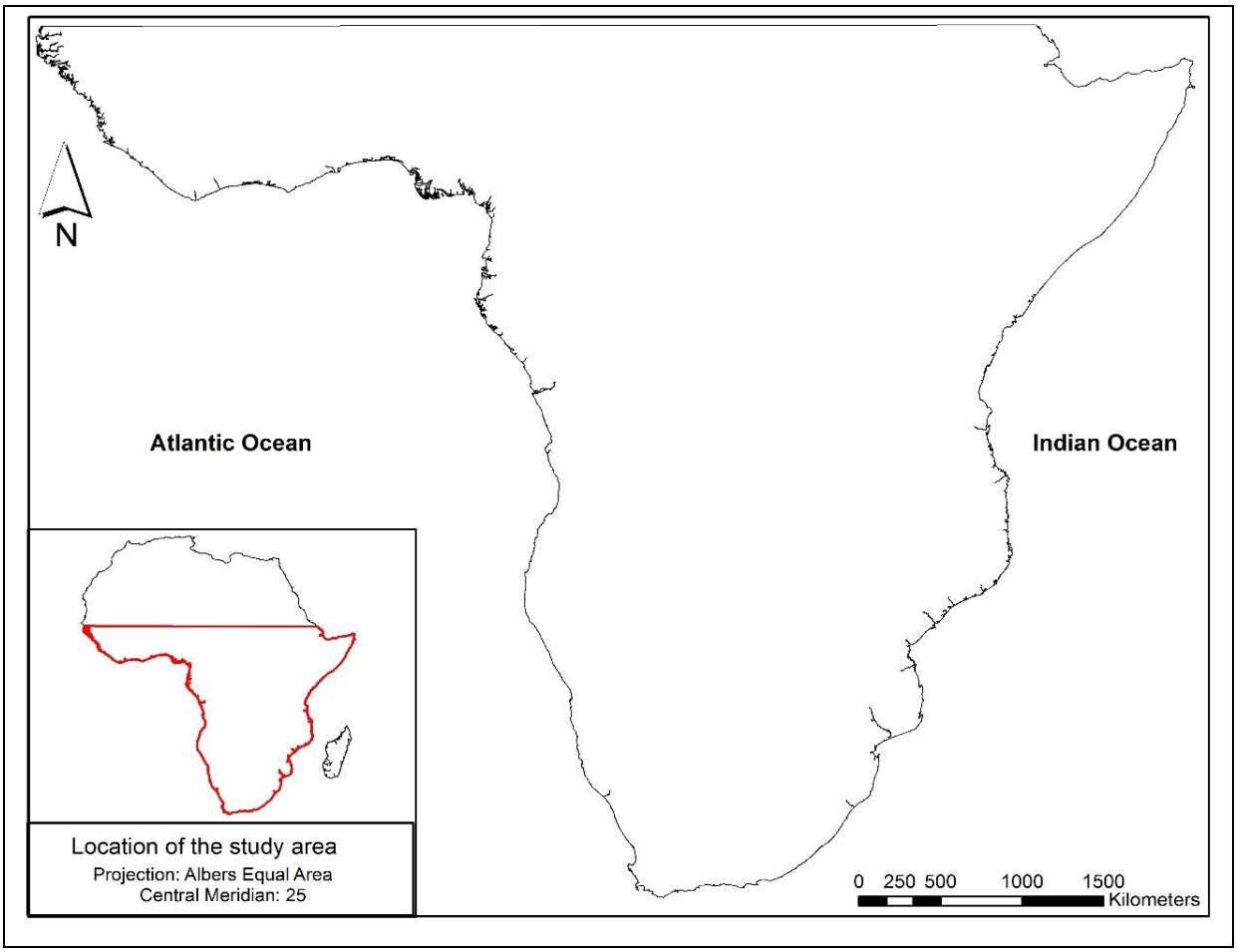


Figure 1.3: Location of the study area with respect to the African continent.

The extent passes through east, west, and central Africa to Cape Agulhas at roughly 35°S latitude and excludes areas north of the Sahel and all offshore islands. The region is bounded by the Sahara Desert to the north, the Indian Ocean to the east through south, and the Atlantic Ocean to the west and covers an area of approximately 18 million km^2.

1.6.1 **Description**

Recognised as one of the World Wildlife Fund's (WWF 2000) eight ecozone's – the Afrotropics constitute one of the most diverse realms in the world, embracing over 500 terrestrial ecoregions. To its northern extreme lies the Sahel, a stretch of semi-arid grasslands denoting the transitional area between the hyper-arid Sahara Desert to the north and undulating subtropical grasslands to the south. In the west, from Sierra Leone to Ghana, lies the Guinean savanna-forest complex, a relatively narrow stretch of evergreen monsoon

forests and grasslands hugged by the warm Guinea current in Atlantic Ocean to its south (Le Houérou 2009). The central extent compromises the Congo Basin, the second largest expanse of uninterrupted equatorial rainforest in the world (Chapin 1923). Drained by the Congo River, mosaics of tributaries, lush perennial forests, savannas, and swamps characterise the region which, owing to its size and severe political tension, remains exceptionally unexplored. The north-eastern Afrotropics are characterised by a vast span of rugged alpine slopes known as the Ethiopian Highlands. These disintegrate into an expanse of lowland subtropical grassland and bushland thickets to the far south-east, known as the African Savanna, home to the largest diversity of megafauna worldwide (Linder et al. 2012.) The sand dune rich Namib and succulent rich Kalahari deserts demarcate the south-west most portion of extent. A biodiversity hotspot, the exceptionally flora-rich Cape floristic region, denotes the southern most extreme of the study extent (Le Houérou 2009).

1.6.2 **Climate**

Climate across the Afrotropics is complex and variable with mean annual temperature ranging from 6˚C, near the alpine slopes of the East African rift valley, to roughly 30 ˚C in the tropical Congo Basin (Figure 1.4a). The distribution of precipitation also ranges widely, from 10 mm per annum in the central Namib desert near the Namibian coast in the south-east to in excess of 10 000 mm near the Cameroonian Highlands which constitutes one of the five wettest localities in the world (Figure 1.4b).

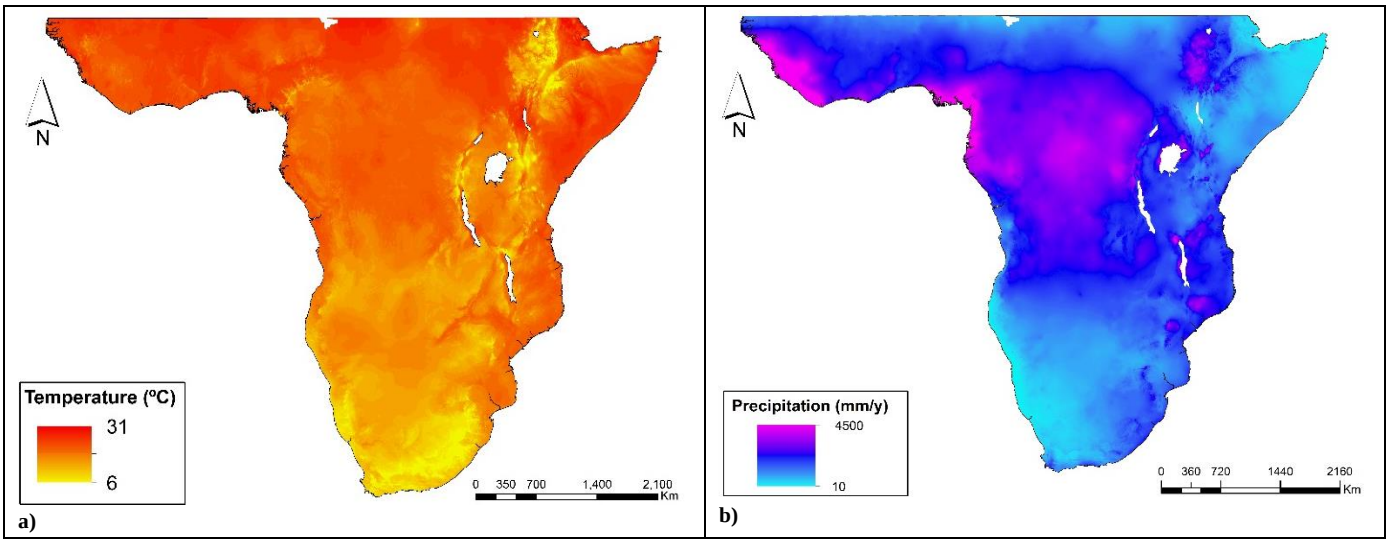


Figure 1.4: Variation in (a) mean annual temperature and (b) mean annual precipitation across the study extent.

1.6.3 **Physiology**

Topography across the extent is also highly heterogeneous (Figure 1.5) with an elevation range of ± 6300m (5800m – -500m). The rugged, undulating slopes of the Ethiopian Highlands and Eastern Arc Mountains comprise the topographically most complex region of the study extent and frequently exceed heights of 4000m while elevation falls to 155m below sea level in the nearby East African Rift Valley.

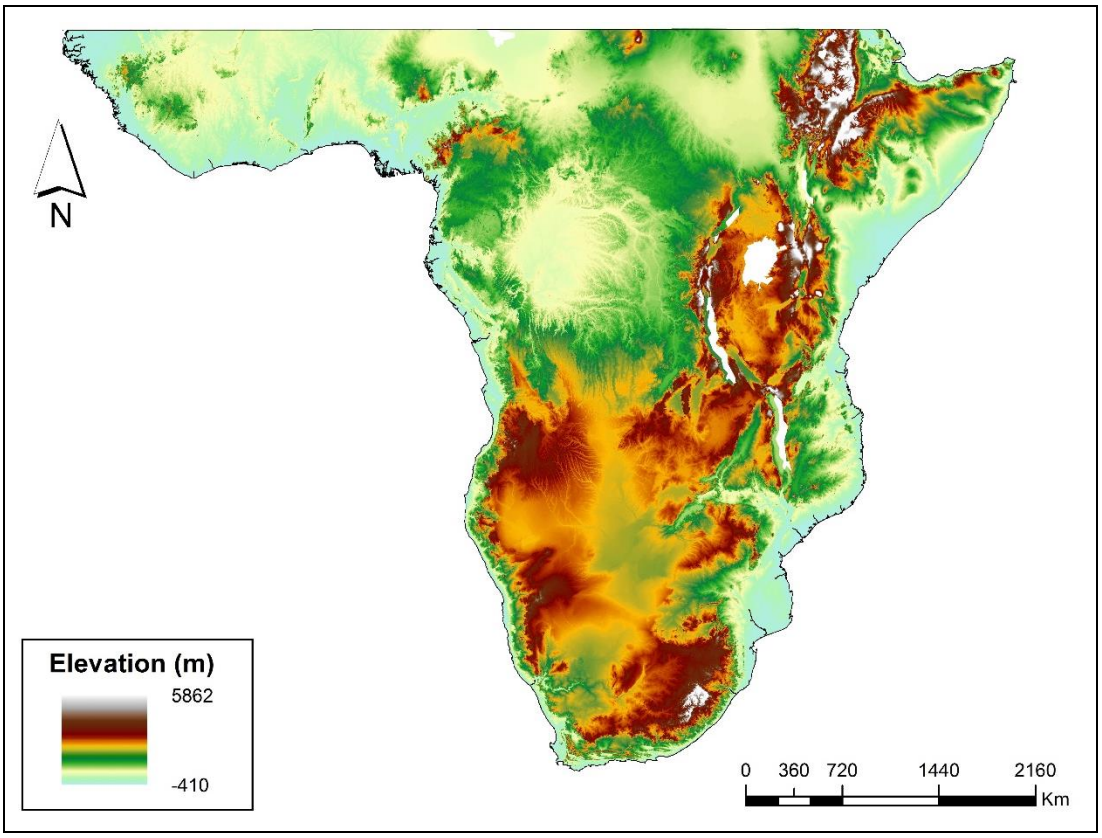


Figure 1.5: Variation in mean elevation across the study extent.

The complex and dynamic climatic and topographic gradients of the Afrotropics, distributed variably across its extent, have sculpted a highly heterogeneous environment which presents an ideal opportunity to examine the influence of environmental processes on patterns of anuran diversity and how these relationships hold across scale.

1.7 THESIS STRUCTURE AND OUTLINE

The thesis comprises five chapters. The current chapter, **Chapter 1**, serves as a prelude for the entire document. It presents an introduction to the study, posits rationale for the research, describes the study extent and provides an overview of the implemented methodology and addresses the overarching aims and objects.

Chapter 2 comprises a literature review. Here, literature relevant to the study is presented and discussed with the objective to familiarise the researcher and reader with important concepts, terms, jargon and methodologies pertinent to the study.

The intermediate section of the document comprises two distinct but interlinked analytical components presented as chapters. Each chapter explores a particular facet of the species-environment relationship and contains its own research focus. In particular, they examine species responses to several real world phenomena using information theory and GIS analysis.

The first research component, **Chapter 3**, examines the relationship between species diversity and the physical environment. In particular, it looks at patterns of species richness, beta turnover and how environmental constituents have differentially influenced the spatial distribution of these.

In **Chapter 4**, the analysis of component one is replicated, but across multiple spatial scales. The objective here is to examine the influence of the modifiable areal unit problem (MAUP) and spatial scale, if any, on the observed patterns and diversity-environment relationships established in component 1.

Chapter 5 concludes the document by revisiting the initial aims and objects of the research then correlating these with the pertinent findings of the study. Limitations encountered by the researcher and recommendations for future research are also presented.

CHAPTER 2: LITERATURE REVIEW

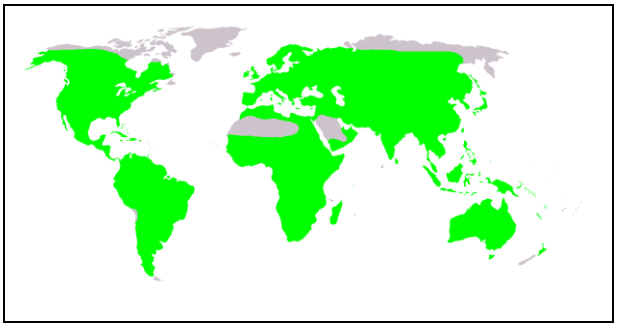
In reviewing existing literature, pertaining to the study, the focus was to:

1. Gain an understanding of amphibian biogeography and correlates of amphibian diversity.
2. Gain an understanding of components of species diversity, including how to measure these.
3. Gain conceptual understanding of the spatial scale and the Modifiable Areal Unit Problem (MAUP) and its relationship with species diversity.

2.1 ANURAN BIOGEOGPRAHY AND CONSERVATION STATUS

Frogs and toads are a highly diverse and largely carnivorous group of vertebrates comprising the order Anura with class Amphibia (Shaffer et al. 2000). Physically they are characterised by their soft, water permeable skin, short front and long hind legs that fold underneath them. Anurans are ectothermic, requiring sufficient amounts of ambient energy for automotive, metabolic and thermoregulatory purposes (Alford & Richards 1999; Carey et al. 2001; Whitton et al. 2012). Their ecothermy allows them to exploit energy poor resources and occupy a variety of niches.

Anurans are widely distributed, being present on all continents, with the exception of some off shore islands and the poles (Figure 2.1). Of the almost 5000 species identified globally, most tend to inhabit warm, moist, aquatic habitats as they require heat for energy and water for reproduction (Buckley & Jetz 2008). Anuran diversity levels are generally highest in the tropics (Roy 1997).



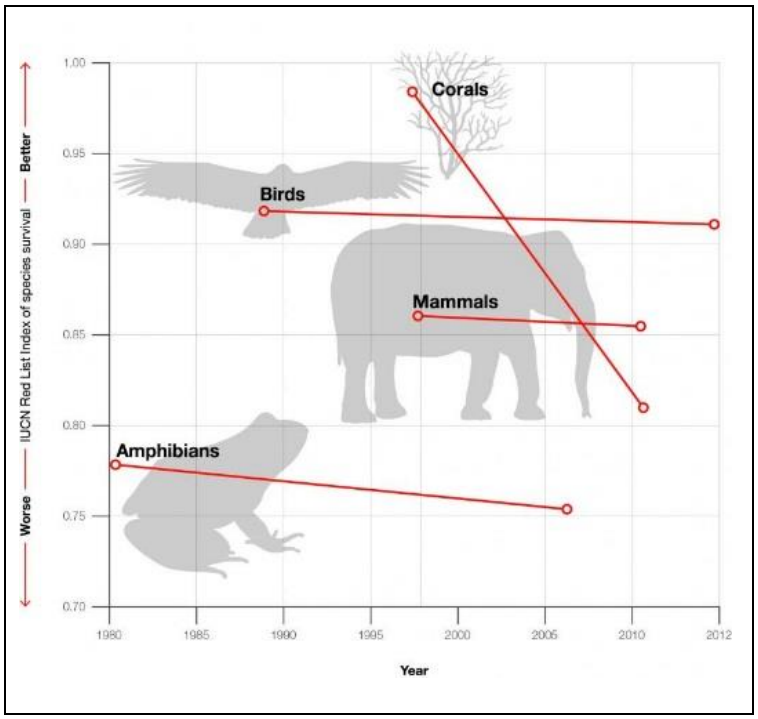
Source: Roy (1997)

Figure 2.1: Global distribution of anurans.

Anurans occupy narrow geographical ranges which are tightly constrained by their poor vagility and acute sensitivity to environmental conditions (Shaffer et al. 2000). The latter trait

has led community ecologists to consider anurans as indicators of broader ecosystem health, often employing them as surrogates in studies pertaining to environmental and climate change modelling (Blaustein & Johnson 2003; Blaustein, Romansic & Kiesecker 2003).

As with most biodiversity, anuran populations are in a precipitous state of decline worldwide with the WWF reporting the extinction of some 170 species the past quarter century (Figure 2.2)



Source: IUCN (2014)

Figure 2.2: Global decline in amphibian populations the past 20 years.

Similarly, the International Union for the Conservation of Nature (IUCN) classifies 10% of all identified anuran species as 'endangered[1]' and some 40% as 'vulnerable[2]' – more than any other order of terrestrial vertebrate (IUCN 2014). Stress-induced pathogenesis, climate change, habitat modification and destruction, particularly through anthropogenic activity, are cited as primary causes for the declines (Alford & Richards 1999; Houlahan et al. 2000; Stuart 2004).

[1] "Species which are likely to become extinct, with few individuals surviving in the wild"

[2] "Species which are likely to become extinct unless current threats are removed" (Global Species Assessment, IUCN 2014)

2.2 DIVERSITY

2.2.1 Background: Whittaker, species diversity and species turnover

A seminal study related to understanding species transitions was conducted in 1960 by renowned ecologist, R.J Whittaker (Whittaker 1960, 1972).Whittaker investigated the variation in local plant species composition in the Siskiyou Mountain forests of the Western United States over several environmental gradients, such as climate and topography. The study revealed that plant species composition , the number of species occupying a site, varied considerably across both topographic and climatic gradients (Whittaker 1960, 1972). The strong association observed between species composition and relative position in an environmental gradient, led Whittaker to the conceptualisation of the term "gradient analysis". Gradient analysis has been defined as a quantitative analysis related to the change in species composition across an environmental gradient usually assessed by means of ordination techniques (Gosz 1992). It provides insight into how species populations, community characteristics and structure, change in response, to or in occurrence with, gradients of the environment. Gradient analysis focuses primarily on changes in species composition within a single habitat along an environmental gradient (Whittaker 1972).To encompass changes in species diversity between different habitats, Whittaker coined the term *beta diversity* (β-diversity).

Whittaker (1960) defined beta, or between habitat diversity, as "the extent of change of community composition". Conceptually, beta diversity is composed of two key components or levels namely, alpha and gamma diversity. Alpha diversity refers to the local diversity or diversity within a single stand, while gamma diversity refers to regional diversity or the diversity of a number of community samples (Jurasinski, Retzer & Beierkuhknlein 2009). The empirical nature of these concepts is elaborated on in section 2.4.

Furthermore, beta diversity is often interchangeably used with, though considered a coarser form of, the concept of species turnover (Louda 1999; Maxime et al. 2008). Korhonen (2014) defines species turnover as the "the rate at which an index of assemblage similarity declines with time" while Vellend (2001) defines it as "the magnitude of change in species composition, along a predefined spatial or environmental gradient". Additionally, Buckley & Jetz (2008) as "changes in species composition along spatial and environmental gradients", respectively. Conceptually, turnover is multifaceted in nature comprising both an ecological and spatial component, which makes it difficult to describe with a single definition. The

concept of turnover thus has both an ecological and spatial component and hence, is biogeographical in nature (Dobrovolski et al. 2012; Calderón-patrón et al. 2013). Here, species composition may refer to species richness, biomass or the number of species per unit area. Turnover can be considered as an indicator of relative change in compositional biodiversity and is of critical importance in understanding the underlying patterns and processes (Vellend 2001) that govern the distribution of species at both regional and global scales (Mac et al. 2004; Burkle, Myers & Belote 2016).

In to order understand why suites of species change in composition or terminate their range edges the confounding factors that govern their distribution, composition and abundance must first be discussed and considered. The next section explores these.

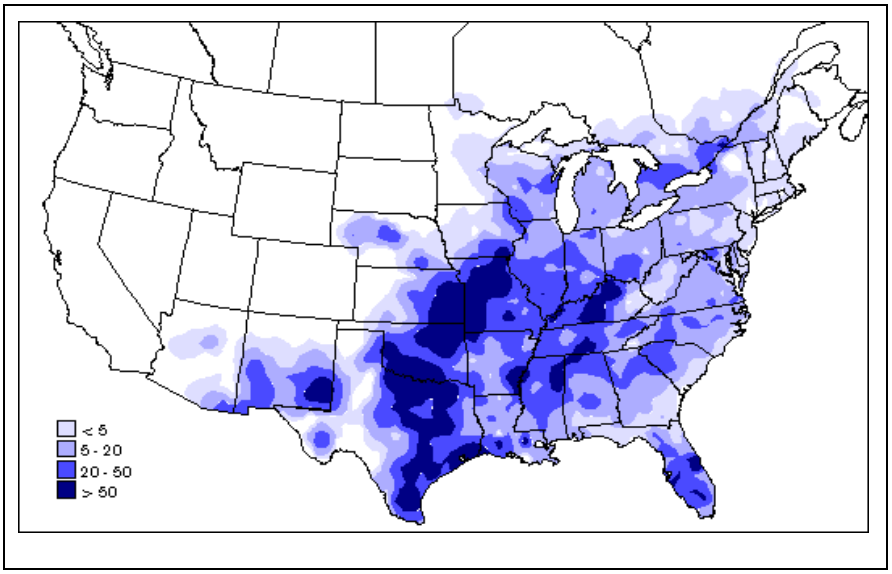
2.2.2 **The nature of ecological communities**

The basic problem of ecology is to determine the causes of the distribution and abundance of organisms (Brown & Kodric-Brown 1977). Every organism or species lives in a matrix of space and time. The space the organism finds itself in is referred to as its habitat, which is determined by its range of tolerance to both biotic and abiotic factors (Tokeshi 1993). Species that share the same habitat, and hence are adapted to similar environmental conditions, form an ecological community. The combination and number of species that inhabit a community, define the community.

The spatial arrangements of habitats define and confine the distribution of a species or species assemblages and communities. Species distribution can be defined as “the manner in which a species is spatially arranged” or “the range size and spatial structure of a species” (Guisan & Thuiller 2005). Simply stated, species distribution is associated with both the geographic range and spatial arrangement of individuals in a local population (Hui 2009). Most distribution patterns of species are dictated by environmental conditions which can change seasonally, in response to both the availability of resources and observational scale of study (Guisan & Thuiller 2005). The concept of species distribution is inherently related to both species abundance, defined as “the number of individuals belonging to a species found per unit sample” (Volkov et al. 2003) while relative species abundance defined as “the evenness of distribution of individuals among species in a community” (Magurran 2004). Species distribution is considered a facet of species abundance.

In return, species abundance is related to and often interchangeable with the concept of species density (Gaston 2000). Density and abundance of a species tends to be greatest near

the centre of its range, where conditions are optimal, and declines gradually toward the extreme boundaries (Guisan & Thuiller 2005; Buckley & Jetz 2007). This rule or principle is applicable to most, but not all species and holds both within steep geographically restricted gradients of environmental change and over entire geographic ranges (Rahbek 2006). Guisan & Thuiller (2005) accredits this pattern to the combination of biotic and abiotic variables that are required for survival and reproduction of individuals of a species to be more ideal towards the central habitat. These requirements define the dimensions of Hutchinson's niche theory which is discussed in later sections. Guisan & Thuiller (2005) further states that the decay in the number of individual of species towards their range edge can be attributed to progressively dissimilar environmental conditions from the central optimal site, resulting in species niche requirements being met less frequently and hence a decrease in its population numbers. Species distribution and abundance patterns across geographic space are depicted by species range and abundance maps (Figure 2.3).



Source: Thompson, Anderson & Bartlein (1974)

Figure 2.3: The relationship between species distribution and abundance. The distribution of a species is depicted by the distribution boundaries of the blue area, while the shade intensity is proportionate to the abundance or density of individuals per unit area.

Species distribution and species abundance, in particular relative species abundance, are both related to the concept of species richness which refers to the total number of species present in an ecological community or landscape. Species richness and relative abundance together determine the species diversity of an ecological community (Peet 1974; Réjou-méchain & Hardy 2012)

For example, a community housing 50 species would be more diverse than one with only 5. When communities contain the same amount of species, the 'evenness' or relative abundance of each species will determine which community is ecologically more diverse. A community

with skewed evenness is less diverse than one where evenness is relatively constant amongst species (Tokeshi 1993).

Species distribution, abundance, richness and composition are largely governed by the same abiotic and biotic variables (Magurran & Henderson 2003). Understanding the implications of each variable is needed to not only address why, where and how species terminate their range but also aids in understanding where changes in composition and subsequent transitions are likely to occur. These in return are important for the establishment and mapping of ecotones, turnover zones and overall biodiversity. In the next section, we look at several variables or factors hypothesised to govern the distribution of amphibian biodiversity.

2.3 FACTORS INFLUENCING AMPHIBIAN DISTRIBUTIONS

Processes generating spatial patterns or the distribution of species can be either exogenous or endogenous (Fortin & Dale 2005). In the case of exogenous pattern generation, the spatial pattern is generated by factors independent of the variable of interest. Exogenous processes include environmental gradients, geomorphological processes and climatic constraints which can cause species distributions to be spatially dependent. Exogenous processes can act independently or interactively of each other. In the case of interactive exogenous processes, the interaction can either be additive when the interactions are linear or multiplicative when the interactions are non-linear (Anselin 2002; Kissling & Carl 2008; Beale et al. 2010). In the case of endogenic pattern generation, the observed spatial pattern is an inherent property of the variable of interest itself (Dole, Doledec & Statzner 2012). Endogenous processes include dispersal capacity, biotic interactions, evolutionary processes and spatial inhibition. Endogenous processes may be entirely inherent or reflect the evolutionary response of taxa to an exogenous process.

As this study focuses primarily on exogenous pattern generation of amphibians through environmental processes, the review of literature focuses specifically on such variables and hypotheses. However, an overview of endogenic processes is also provided because endogenic and exogenic factors are not mutually exclusive (O'Neill et al. 1991; Cornell & Lawton 1992).

2.3.1 **Endogenic factors**

2.3.1.1 Ecological interactions

An ecological community is an assemblage of species in the same place or habitat. Interactions between species can either cause harm or provide benefits for the organisms involved. In return such interactions affect the survival and reproduction of individuals and hence the local diversity of a community (Cornell & Lawton 1992; Burkle, Myers & Belote 2016). Table 2.1 provides an overview of the most prominent species interactions known to affect local diversity and the corresponding consequences of these interactions.

Table 2.1: Species interactions and the effects thereof on diversity as adapted from Cornell & Lawton (1992).

Type of Interaction	Effect on diversity
Interaction	Effects on interacting populations
Predation	Prey disperse from areas where predation is high
Herbivory	Where herbivory is high plant species diversity decays
Competition	Competing species lose access to some resource; this can result in speciation and niche modification and hence increase in diversity or a decrease in diversity via the local extinction of a species.

2.3.1.2 Dispersal limitation

Ronce (2007) defines dispersal as "the movement of individuals or species with potential consequences for gene flow across space". Dispersal holds a central role for both the dynamics and evolution of spatially structured populations, thus allowing the genetic cohesion of a species across space, its global persistence despite local extinction, and the tracking of favourable environmental conditions. Species with lower environmental tolerances generally have poor dispersal capacity while those with broader environmental tolerances have greater dispersal capacity (Mac Nally et al. 2004; Rouquette et al. 2013). Beta diversity is expected to be higher in species with low dispersal capacity and lower in species with high dispersal capacity (Wang et al. 2012; Calderón-patrón et al. 2013).

2.3.1.3 History

Time, in an ecological and evolutionary sense, is one of the most important determinants of species distribution and richness (Wiens & Donoghue 2004). More time permits for the colonization of new habitats, physiological adaptations, and the evolution of more species via speciation. Subsequently, ecologically older realms which are biographically similar, tend to have higher levels of species diversity (Wiens & Donoghue 2004; Fjeldsa & Rahbek 2006).

2.3.2 **Exogenous factors: Environmental factors affecting amphibian diversity**

2.3.2.1 Water availability

The distribution of terrestrial biodiversity has been positively linked to moisture availability (Currie 1991; McCain 2007; Chen et al. 2011). This relationship has been documented to persist for a variety of taxa across a range of spatial scales and biogeographic zones (Kiesecker, Blaustein & Belden 2001; Sodhi et al. 2008; Ortiz-Yusty, Páez & Zapata 2013; Poynton 2013). Ortiz-Yusty, Páez & Zapata (2013) studied the relationship between amphibian diversity and annual precipitation in Columbia and found amphibian diversity to be highest in moisture saturated areas. In a meta-analysis of studies analysing environmental factors influencing global vertebrate diversity, Hawkins et al. (2003) found precipitation to correlate best with amphibian diversity.

Being predominantly aquatic, anurans require sufficient amounts of moisture for both survival and reproduction (Buckley & Jetz 2007, 2008). Water availability thus comprises a salient variable in determining the spatial distributions of anurans.

2.3.2.2 Temperature

Currie (1991, p.28) postulates in his 'species-temperature' hypothesis that "benign conditions permit for more species" and that regional species diversity is mediated by the thermoregulatory need of that species. Simply, he suggests that temperature, or ambient energy, acts as a physiological constraint - a threshold of sorts - above or below which a species cannot perform sufficiently (Zuo et al. 2012; Ortiz-Yusty, Páez & Zapata 2013). For ectotherms, such as anurans, temperature has been shown to influence rates of energy use, the assimilation of resources, interactions with other species and species mobility (Buckley & Jetz 2007, 2008; Qian et al. 2007; Sodhi et al. 2008; Diniz-filho et al. 2010; Dobrovolski et al. 2012; Martin & Ferrer 2015).

Studying the influence of amphibian diversity in the Andes Mountains of South America, Ortiz-Yusty, Páez & Zapata (2013) found that local fluctuations in diversity correlated significantly with variation in mean annual temperature which they attributed to the ecothermic physiology of amphibians. Similarly, Chejanovski & Wiens (2014) found regional variation in the species richness of North American tree frogs to correlate with mean annual temperature, although the relationship was established as humpback and modulated by precipitation. In a similar study, but for beta diversity in Australia, Guerin, Biffin & Lowe (2013) found that variation in mean annual temperature could explain over 92% of the

variation in faunal turnover, which are higly coupled with amphibian turnover (*see* Buchanan et al. 2008).

2.3.2.3 Environmental energy

The species-energy hypothesis is a climatically based hypothesis which postulates that regional species diversity is mediated by energy availability (Rodríguez, Alfonso & Hawkins 2005; Qian et al. 2007; Qian 2009). It predicts that greater energy availability allows for an increased number of individuals, stable populations and more species to coexist (Evans, Greenwood & Gaston 2005).

Two mechanisms have been proposed for the species-energy relationship (Hawkins et al. 2003). The first suggests that energy determines diversity through its influence on and propagation through trophic cascades (Phillips, Hansen & Flather 2008; McGlynn, Weiser & Dunn 2010; Whitton et al. 2012). That is, more energy, allows for more resources to be used by producers, which in turn allows for more consumers. The second mechanism is nested in niche theory and suggests that energy determines diversity gradients through its influence on species physiology i.e. resources partitioning (Martin & Ferrer 2015).

Examples of energy variables include actual evapotranspiration (AET) and potential evapotranspiration (PET). AET measures the amount of water lost to the atmosphere through both evaporation and transpiration and is an indirect measure of the water-temperature balance (Colwell 1988; Clarke & Gaston 2006; Buckley & Jetz 2007). Conversely, potential evapotranspiration measures the amount of water that would be lost to the atmosphere if sufficient supplies would be available. PET has widely been used as measure of both the water-temperature balance (Buckley & Jetz 2007) and ambient energy (Veech & Crist 2007) in species distribution modelling. For example, Hawkins et al. (2003) linked global vertebrate diversity to energy availability. Specifically, that higher vertebrate diversity correlated significantly with higher energy availability as measured by AET and PET. Similarly, Rodríguez, Alfonso & Hawkins (2005) coupled herpetological diversity across Europe to PET while Qian et al. (2007) found amphibian diversity to correlate highly with variation in AET across continental China.

2.3.2.4 Environmental productivity

Productivity is a measure of ecosystem energy (Clarke & Gaston 2006). Productive-energy metrics record the amount of resources available for consumers to turn into biomass. The ‘more individuals hypotheses’ (MIH) links productive energy to diversity and postulates that

productivity drives diversity by lowering extinction rates owing to larger population sizes (Currie 1991; McGlynn, Weiser & Dunn 2010). This implies that more productive resources permit for more species. Additionally, McCain & Grytnes (2010) suggests that productive environments may promote diversity through niche conservatism and adaptations.

Remotely sensed proxies, including the Normalised Difference Vegetation Index (NDVI) and net primary productivity (NPP), are commonly used as measures of ecosystem productivity in species distribution modelling. The NDVI is an indicator of chlorophyll content and photosynthetic activity derived through the manipulation of spectral bands. Net primary productivity (NPP) is a measure of photosynthetic energy retained by plants and derived from the NDVI.

Numerous studies have linked vertebrate (Pettorelli et al. 2005; Rahbek 2005; Qian 2010) and amphibian diversity (Werner et al. 2007; Whitton et al. 2012) to ecosystem productivity. Results between studies have not been equivocal with some finding the relationship between productivity and diversity as either negative (Gaston et al. 2007) or positive (Lausch et al. 2013). The relationship thus seems to be taxa and region specific. For example, Martin & Ferrer (2015) positively correlated amphibian turnover to variation in seasonal NDVI across the Mediterranean basin while Gaston et al. (2007) found avian turnover to decrease with increasing productivity. Irrespective of the nature of the relationship, productivity comprises a salient abiotic component of ecosystem diversity.

2.3.2.5 Climatic stability

The environmental stability hypothesis posits that stable environmental conditions promote for higher species diversity due to a constant supply of resources promoting the conservatism of some niches and the evolutionary adaptations of others (specialisation). Whitton et al. (2012) linked amphibian richness and range size to climatic variability and found that areas of high climatic stability often had higher species richness and smaller ranges, which they attributed to niche breadths and lower environmental tolerances of species associated with such regions.

2.3.2.6 Climatic extremes

The climate extreme hypothesis, intrinsically linked to the stability hypothesis, postulates that species distributions are constrained by climatic extremes. In particular, it predicts that areas with greater climatic extremes have lower species diversity because such areas select for generalist, more wide ranged species which possess broad climatic niches (Diniz-filho et al.

2010). In contrast, areas associated with limited climatic extremes selects for small ranged, specialised species with low environmental tolerances, narrow niche breadths and limited dispersal capacity.

In a study correlating the global distribution of amphibian ranges sizes to various climatic hypothesis, Whitton et al. (2012) linked amphibian range size to extremes in temperature ranges. Studying North American tree frog richness, Chejanovski & Wiens (2014) found that areas of higher climatic stability and limited climatic extremes had higher species richness which they attributed to speciation events.

2.3.2.7 Physiography

Topography is not an explicit environmental process but may influence patterns of diversity patterns in two ways. First, steep topographic gradients may drive gradients of diversity by acting as dispersal inhibitors (Adler & Lauenroth 2003; Wang et al. 2012) to species thereby influencing species richness and beta diversity through the promotion of speciation. Second, altitudinal variability causes variation in climatic and hence biogeographic conditions over short distances which too may drive gradients of diversity particularly at a mesoscale by broadening niche breaths (McCain 2007). Amphibians possess poor dispersal capacity (Pineda & Lobo 2009) and thus dispersal barriers such as mountains may promote diversity, both richness an turnover, through speciation and resource partitioning. Numerous studies have linked the spatial turnover of amphibians to dispersal inhibitors such as elevation gradients (Qian & Shimono 2012; Tang et al. 2012; Bishop et al. 2015; Murphy et al. 2015) with higher turnover in areas with greater elevation gradients.

2.3.2.8 Habitat heterogeneity

If the rationale of niche theory is considered then habitats that are physically or biologically complex will furnish more niches and thus house a wider variety of species (Veech & Crist 2007). Transitional zones between two different biomes or biogeographical areas – called ecotones – tend to be highly diverse because they include species from both neighbouring biomes as well as species unique to the transitional area. The relationship between habitat heterogeneity and species diversity has not been equivocal. Studying the species diversity of avian montane assemblages, Veech & Crist (2007) found that high habitat heterogeneity correlated positively with gamma diversity negatively with both alpha and beta diversity. Similarly, Silva et al. (2011) found that anuran alpha diversity was higher in environmentally

homogenous areas and lower in heterogeneous areas. In contrast, Katayama et al. (2014) found species richness was highest in areas of high habitat heterogeneity.

2.4 MEASURING DIVERSITY

There are many ways to measure biological diversity (*see* Magurran 2004). In an ecological context, biodiversity refers to the different types of species and their relative abundance or evenness at a given spatial scale (Colwell 1988). Diversity is traditionally described by means of two components: inventory diversity, which describes the number of species within a sampling quadrat (α) or across a region (γ) and differentiation diversity or beta diversity (β), the change in species composition between regions (Diserud & Ødegaard 2007). Inventory and differentiation diversity are inherently related but are measured in different ways. In the succeeding sections, an overview of each concept and associated measures are described and supplied.

2.4.1 Inventory diversity: Alpha and gamma diversity

Inventory diversity comprises two components, local diversity and regional diversity. Local or within-habitat diversity is termed *alpha diversity* (α) while regional or within-landscape diversity is termed *gamma diversity* (γ) (Loreau 2000; Jost 2007; Tuomisto 2010a; Zhang et al. 2014). Both alpha and gamma diversity are measured in sampling plots or quadrants (Figure 2.4), the size of which can range from a few meters (local scale) to thousands of square kilometres (regional and continental scales) (Veech & Crist 2007).

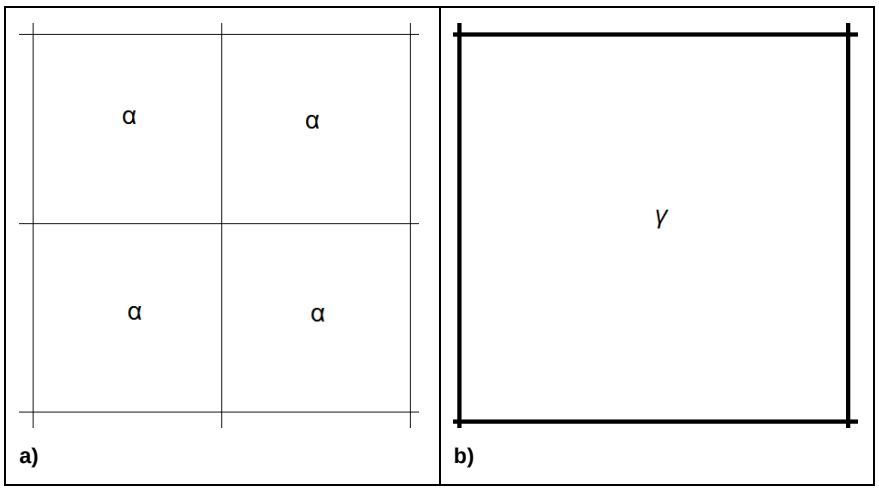


Figure 2.4: (a) Alpha diversity and (b) gamma diversity measured using sampling quadrats

Both diversity components measure the same aspect of community diversity (i.e. richness, abundance, or both) and are differentiated only by the scale at which they are measured

(Ulrich & Almeida-neto 2012). For all purposes, only measures of alpha diversity are discussed as these fall central to the scope of the study.

2.4.1.1 Measures of alpha diversity

Alpha diversity can be measured in a variety of ways, the simplest of which is to count the number of species in a sampling unit. Initially, Whittaker (1960, 1972) proposed that Fisher's alpha parameter of the log series be used to measure alpha diversity (Volkov et al. 2003; Cayuela et al. 2006). Since then however, numerous indices have been incorporated into ecological literature to measure alpha diversity.

Apart from simply counting the number of species in a sampling unit, measures of alpha diversity can be partitioned into two primary groups: species richness estimators and diversity indices. Estimators of species richness approximate the total number of species present in a community or sampling quadrat and include species-accumulation and rarefaction curves (Jurasinski et al. 2012). In contrast, diversity indices, referred to as heterogeneity indices, combine both species richness and abundance into a single value of evenness. Communities that are numerically dominated by one or a few species exhibit low evenness while communities where abundance is distributed equally amongst species exhibit high evenness (Reedk et al. 1996). Citing all measures of alpha diversity falls beyond the scope of this literature review; for this reason, an overview of some commonly used richness estimators and diversity indices is supplied by Table 2.2.

Table 2.2 Commonly employed indices of alpha diveristy.

Species diversity component	Index	Reference
Richness	Log α	(Colwell 1988)
	Log Normal α	
	Brillouin	
	McIntosh	
	Henderson	
	Shannon	
Abundance	Simpson	(Blackburn & Gaston 2002; Tuomisto 2010b)
	Shannon Evenness	
	Fisher's α	
	McIntosh E	

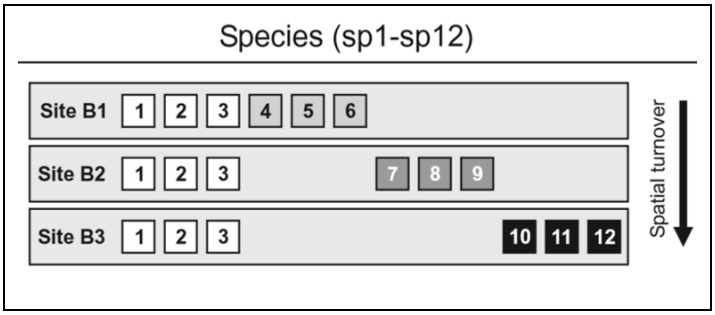
All indices of inventory diversity, are to some degree, sensitive to the size of the sampling unit used to measure diversity, the sampling effort, and habitat heterogeneity (Peet 2015). Abundance-based indices have also shown strong sensitivity to species dominance and evenness.

The selection of the most appropriate index or means to quantify alpha diversity will depend on; the appropriateness of the index for the data; the discriminability of the index; statistical comparability; widespread utilization of the index or measure to ensure comparability between studies, sites or researchers; the ecological question being asked and phenomena being addressed.

2.4.2 **Differentiation diversity: Beta diversity**

Investigating changes in diversity at a given scale is usually done using presence-only, presence-absence, or quantitative species-abundance data (Tuomisto 2010b). Presence-only data comprises a list of the known locations of species while presence-absence data refers to data derived from inferring the known occurrence of individual species and are usually represented by species distribution polygons (i.e. range maps). Presence-absence data considers the dispersal capacity of the individual species. In contrast, species abundance data are quantitative and based on the number of individuals belonging to a species in a sampling quadrat. Species abundance data are limited for large geographic extents owing to the spatio-dynamic nature of taxa and the exhaustive sampling requirement. Because presence-absence datasets are more widely available and cover broader spatial extents than abundance datasets, presence-absence based beta diversity methods are preferred to abundance based ones (Shmida 1984; Ricotta & Pavoine 2015). For this reason, only presence-absence based measures of beta diversity metrics are described here.

Beta diversity can be divided into two components: turnover and nestedness (Williams 1996; Baselga 2010; Almeida-Neto, Frensel & Ulrich 2012; Navarro-sigüenza & Rodríguez 2014). Both concepts have roots in Whittaker's (1960) initial concept of β-diversity, " the extent of change in community composition between sites" (Tumoisto & Ruokolainen 2008), but can conceptually be distinguished from one another. Species turnover refers to the replacement (Figure 2.5) of species from one sampling unit to another across spatial, temporal and environmental gradients (Williams 1996; Lennon et al. 2001; Legendre et al. 2005).

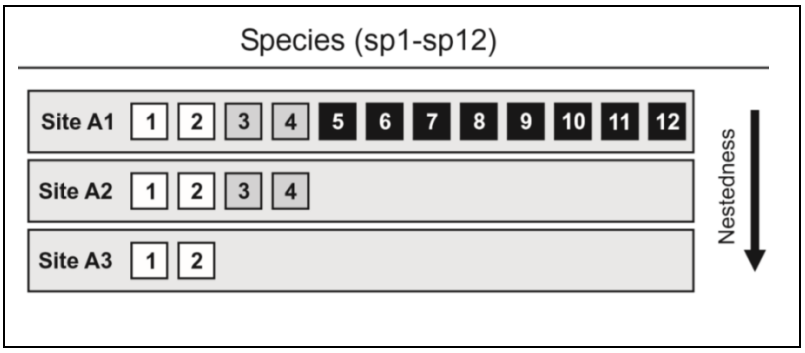


Source: Baselga (2010)

Figure 2.5: Spatial turnover reflecting species replacements between sites.

Here, the changes are attributed to environmental sorting as well as spatial and historical constraints such as dispersal limitation (Soininen, Lennon & Hillebrand 2007). Additionally, turnover is usually expressed as a rate and is associated with defined directional gradient of interest (Anderson et al. 2011).

Nestedness, on the other hand, focuses on the fluctuation in community composition caused by the gains and losses of species between sites or sampling units (Baselga 2010; Dobrovolski et al. 2012). Typically, nestedness occurs when biotas (plant or animal life) of sites with smaller number of species are subsets of biotas at richer sites. Nestedness is considered as the outcome of some non-random spatial process which promotes the disaggregation of assemblages and the loss of species between sites (Figure 2.6).



Source: Baselga (2010)

Figure 2.6: Nestedness reflecting the grains and losses of species between sites.

Turnover and nestedness both comprise aspects of beta diversity "the change in species composition between sites" but are antithetic in their biological consequence (Bishop et al. 2015). Therefore, it is important to distinguish between the two components of beta diversity as they reflect the outcome of different spatial and ecological mechanisms (Figure 2.7).

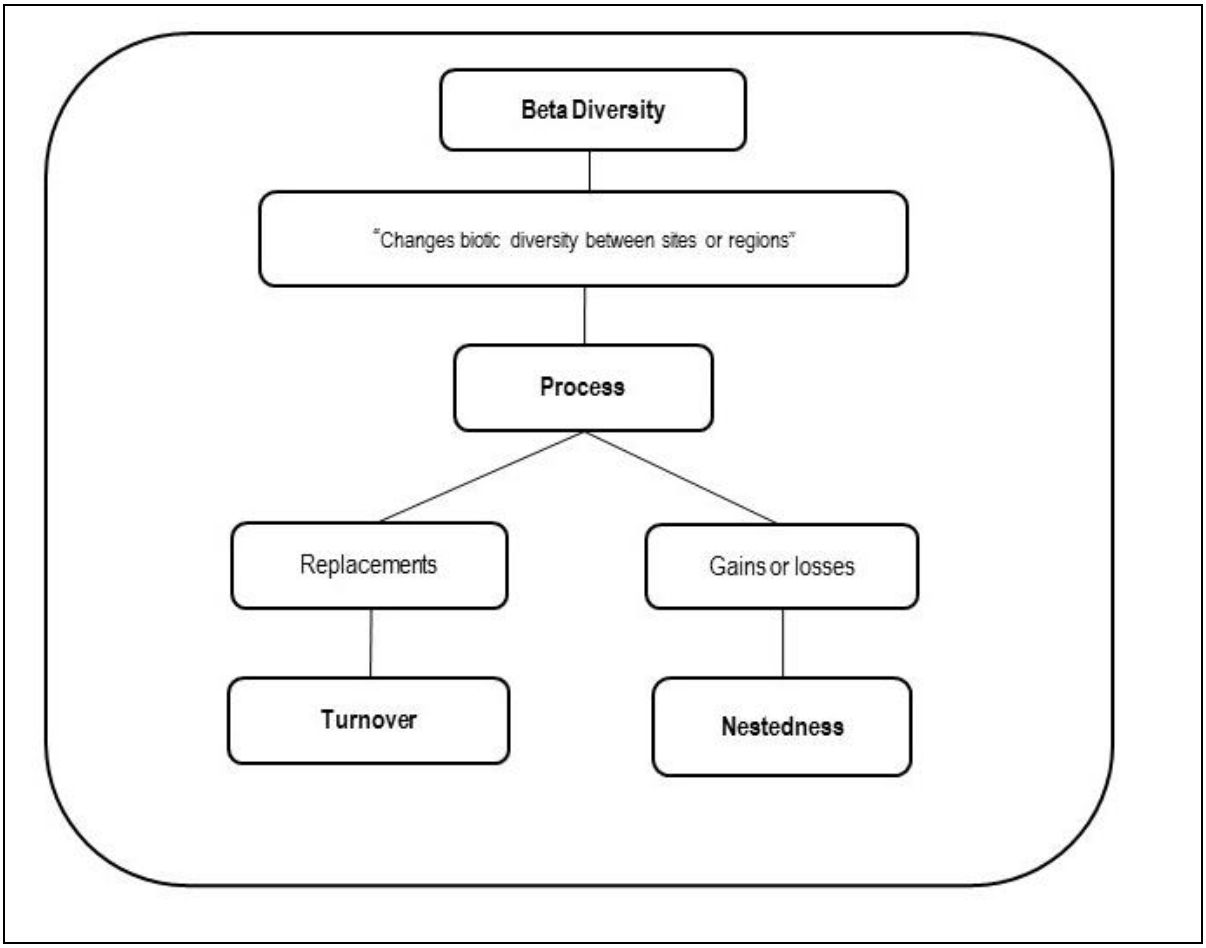


Figure 2.7: Partitioning beta diversity into its turnover and nestedness components.

Various measures of beta diversity as turnover have been introduced (Koleff et al. 2003), however there is no overall consensus as to which are most appropriate for addressing particular ecological problems (Buckley & Jetz 2008; Hui 2009, Hui & McGeoch 2014). Rather, the type of measure employed must reflect the question being asked or ecological phenomena being addressed.

Anderson et al. (2011) state that β-diversity has two essential classes of measurements, which can be used to describe species turnover, namely:

1. Classic metrics or traditional measures of beta diversity expressed using α and γ components.
2. Multivariate metrics expressed using similarity and dissimilarity indices.

In the subsequent sections, descriptions of each class and its associated measures are supplied.

2.4.2.1 Classic metrics of beta diversity

Whittaker's (1960) classic β-diversity concept is composed of two empirical components namely, alpha and gamma diversity (Magurran & Henderson 2003; Tuomisto 2010b). Alpha diversity (α) represents the mean number of species at a sampling location at a local scale and gamma diversity (γ) represents the regional biodiversity of an area or plot at a broader scale (Jurasinski et al. 2009).

Empirically, Whittaker defined β-diversity as the ratio between gamma and alpha diversity (Equation 2.1) (Whittaker 1960).

$$\beta_W = \frac{\gamma}{\bar{\alpha}}$$

Equation 2.1: Whittaker's (1960) true beta diversity index.

Whittaker, called this 'true beta diversity' and arrived at this equation by reasoning that if the average species composition within a set of communities or samples are known, then the total diversity of all samples could be found by multiplying the average diversity with the number the number of communities or samples ($\gamma = \beta \times \alpha$). Pure measures of β-diversity, are thus those which examine the extent of the difference between two or more areas of α-diversity relative to γ-diversity.

The simplest form of species turnover is rendered by Whittaker's absolute species turnover (β_A) formula which aims to quantify how much more diversity the entire dataset has (i.e. γ – diversity) than an average composing subunit diversity (Equation 2.2).

$$\beta_A = (\alpha_1 - c) + (\alpha_2 - c)$$

Equation 2.2: Whittaker's (1972) absolute species turnover formula.

Where β_A is the absolute species turnover, α_1-α_2 is the species richness of the two sites being compared, and *c* is the species common between both sites.

If the absolute species turnover (β_A) is divided by α-diversity, then a measure is obtained that quantifies the number of times the species composition changes completely among subunits across the entire dataset. This normalised measure of β_w is referred to as "Whittaker's species turnover" concept (Equation 2.3).

$$B_{W-1} = \frac{(\gamma - \alpha)}{\alpha} \quad \boldsymbol{OR} \quad \beta_{w-1} = \frac{\gamma}{\alpha} - 1$$

Equation 2.3: Whittaker's species turnover formula.

Following a meta-analysis of 60 published beta diversity studies spanning some 40 years, Koleff et al. (2003) noted that Whittaker's initial beta diversity measures were the most widely used. However, several scholars have criticised Whittaker's turnover measures resulting in numerous modifications (some 20 indices) (*see* Jurasinski et al. 2009; Lande 1996; Tuomisto 2010a; Tuomisto 2010b; Tuomisto 2012). For example, Lande (1996) notes that Whittaker's γ-diversity is subject to overestimates when communities share species. Lande thus suggested and proceeded to include additive species partitioning in the concepts of α, β and γ diversity which he condensed into a single formula (Equation 2.4).

$$B_{Lande} = \gamma - \alpha$$

Equation 2.4: Lande's (1996) additive partitioning of beta diversity.

Additive partitioning treats alpha diversity as the average within sample diversity and beta diversity as average amount of diversity not found in a single random sample where both measures share the same unit of measure (Lande 1996; Jurasinski et al. 2009). Unlike Whittaker's formula, Lande's approach is applicable across scale if a standardised sampling scheme is used. Additionally, his approach differs to that of Whittaker's initial approach in that the relationship between variables is additive rather than multiplicative. Lande's approach is similar to that of MacArthur (1965) but became more established in the ecological literature because MacArthur (1965) failed to express his measure of beta diversity explicitly in terms of α and γ components.

Furthermore, Vellend (2001) notes that Lande's formula too is not without bias. For instance, if many small sampling units are used then low γ-diversity and high β-diversity values will be rendered as beta diversity tends to decrease as sample heterogeneity increases. Though this, as noted by Jurasinski, Retzer & Beierkuhnlein (2009), depends on the number and size of the samples in relation to extent of the study area. Vellend (2001) further argues that neither Whittaker's (1960) nor Lande's (1996) measures truly reflect the concept of species turnover, "the replacement of one species by another along an environmental or spatial gradient of continua", since neither considers the distribution of species on a spatial or environmental scale. Additionally, he postulates that beta diversity, as measured by partitioned α and γ components, is simply a value of diversity mathematically related to alpha and gamma diversity and thus the terms *β-diversity* and *species turnover* are not truly interchangeable concepts. Concluding his meta-analysis, Vellend advocated that beta diversity, as turnover, is better measured using similarity indices. These are indices are discussed in section 2.4.2.2.

2.4.2.2 Multivariate measures: Similarity and dissimilarity indices

These are measures which examine the difference in species composition between sampling units (α-diversity). Specifically, they evaluate the distinctness of assemblages (i.e. co-occurring species in a community) in a given habitat and make a pairwise comparison between sampled alpha diversities across the entire study area (Magurran & Henderson 2003; Jurasinski et al. 2009; Tuomisto 2010b).

The similarity is described between a focal quadrat and an adjacent quadrat using three parameters: a = the number of shared between the two sampling quadrats, b = the number of species in the focal quadrat but not the adjacent quadrat and c = the number of species in the adjacent quadrat but not the sampling quadrat (Figure 2.8) (Mcdonald et al. 2005; Qian 2009; Gueze et al. 2013).

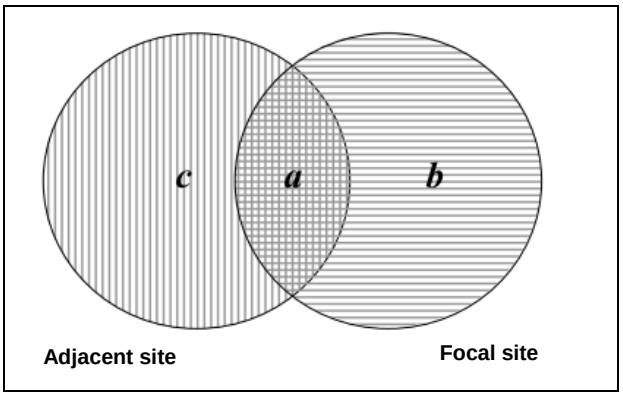


Figure 2.8: Assemblage composition expressed as mismatching components.

These parameters are usually rearranged such that they are ratios of each other and the similarity is standardised on a 0-1 scale, where 1 = complete similarity and 0 = complete dissimilarity. Normalising the similarity index (1 – similarity index) renders a corresponding dissimilarity coefficient. The most commonly used similarity indices include the Sorensen index, the Jaccard index and modified Simpson index (Table 2.3) (Lennon et al. 2001).

Table 2.3: Indices of diveristy used to measure beta diversity as species turnover.

Index	Formulation	Literature
Sorensen	$\beta_{Sor} = \frac{2a}{2a+b+c}$	(Dobrovolski et al. 2012)
Jaccard	$\beta_{Jac} = \frac{a}{a+b+c}$	(Podani & Schmera 2016)
Modified Simpson index	$\beta_{Sim} = \frac{\min(b,c)}{\min(b,c)+a}$	(Lennon et al. 2001)

For example, Vasudevan, Kumar & Chellam (2006) used the Sorensen dissimilarity index to measure the species turnover of amphibians in the mountains of southern India while Di

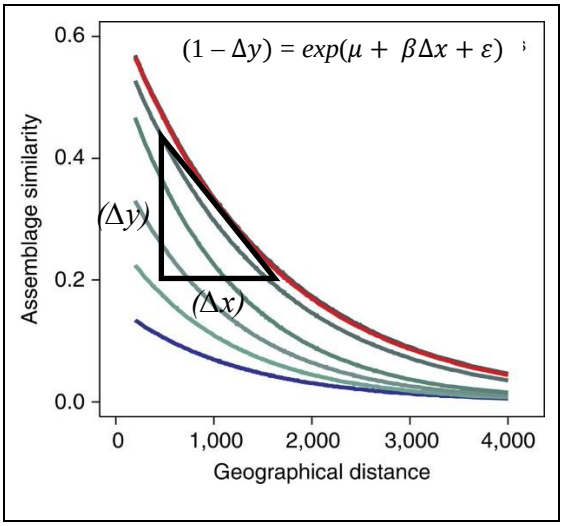
Virgilio, Laffan & Ebach (2012) used the Sorensen index to measure the spatial turnover of mammals along ecotones in New Zealand. Similarly, Qian (2009) used the Jaccard dissimilarity index to measure the turnover of terrestrial vertebrates at a global scale while Mcdonald et al. (2005) used the same index to determine the turnover of avifauna along biographic transitional zones in Canada.

The Sorensen and Jaccard indices have been criticised for being sensitive to species richness gradients, therefore the Simpson dissimilarity index is often preferred to measure turnover as it has been found to be insensitive to richness gradients (Legendre 2014; Ricotta & Pavoine 2015b). Bishop et al. (2015) used the Simpson index to measure and map the spatial turnover of ant assemblages flanking the Drakensberg mountain chain in eastern South Africa while Keil et al. (2012) used the same index to measure and map the spatial distribution of turnover over Mediterranean avian assemblages across mainland Spain. Baselga (2010) argues that the Simpson index is the best index for capturing species replacement while the Sorensen and Jaccard indices are best used to capture the overall beta diversity (compositional (dis)similarity) of a region. Almeida-Neto, Frensel & Ulrich (2012), however, have rejected this conjecture and have found that the Sorensen and Simpson index render the same values when two sites have the same species richness and when there are no shared species between two sites. They opine that the selection of the index used to measure turnover remains arbitrary and dependent on the research aims although they propose and recommend the use of an alternate version of the Simpson index to measure species turnover.

2.4.2.3 Rates of turnover

Species turnover is often expressed as rate. To measure rates of turnover, similarity index values are log transformed to avoid distributional problems and then put as a function of some spatial (i.e. distance-decay) or environmental distance (Buckley & Jetz 2008; Fitzpatrick et al. 2013; Schoener 2015). A model trend is then fit through the data. The rate of change of the fitted trend line (i.e. slope) represents the directional turnover rate. The steeper the slope of the trend line, the more rapid the turnover.

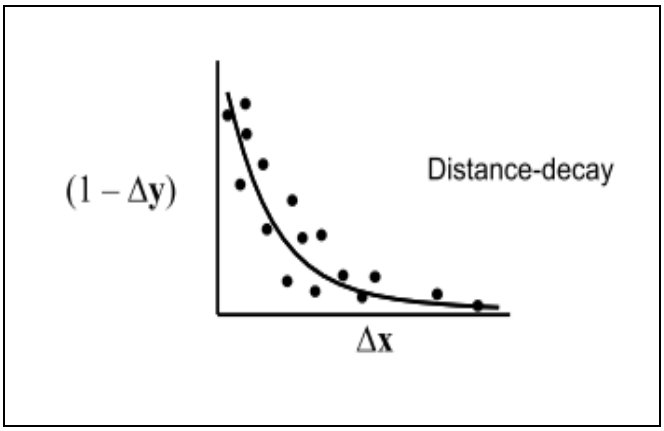
As adapted from Anderson et al. (2011), let y denote some similarity co-efficient and x the distance across samples. Let the relationship between variables on a log scale displays a decaying exponential trend. Beta diversity, as turnover, will be measured as: $(1- \Delta y) = exp(\mu + \beta \Delta x + \varepsilon)$ (Figure 2.9).



Source: Anderson et al. (2011)

Figure 2.9: Hypothetical scenario reflecting how the rate of turnover is measured as adapted from Anderson et al. (2011).

Nekola & White (1999) suggested that species turnover be measured as the slope of the distance-decay relationship curve (Figure 2.10). The reasoning here was that as the distance between sampling units is increased the compositional dissimilarity between those sampling units is increased. The slope of the fitted regression curve represents the rate of compositional turnover across geographic space and thus explicitly relates beta diversity to a gradient as initially envisaged by Whittaker (1960). This approach has become widely adopted in the ecological literature and is often used to test theories and hypothesis pertaining to the influence spatial processes exert on the spatial organisation of communities (Tuomisto 2010a; Qian & Ricklefs 2012; Qian & Shimono 2012; Calderón-Patrón et al. 2013; Saito et al. 2015).



Source: Anderson et al. (2011)

Figure 2.10: Measuring species turnover as the slope of the distance-decay plot.

As illustrated above, various measures of beta diversity exist. Each measure emphasises a different aspect of assemblage structure and therefore yields varying results. Ultimately, the measure selected will depend on questions being asked or ecological phenomena being considered.

2.4.2.4 Beta diversity: New measures

Traditional and multivariate measures of beta diversity are based on pairwise comparisons of individual assemblages only (Anderson et al. 2011; Barton et al. 2013; Barwell et al. 2015). That is to say, when comparisons are made between three or more assemblages, only the average pairwise similarity is used (Hui & McGeoch 2014; Lande 1996). Therefore, the diversity components of several assemblages cannot be fully expressed with only γ and β diversity components. Recently, Hui & Mcgeoch (2014) proposed zeta diversity (ζ), an incidence-based metric that captures all diversity components produced by assemblage partitioning in order to compute beta diversity.

Zeta diversity builds on the limitations of existing multivariate incidence metrics which can only quantify compositional similarity, by means of pairwise comparisons between the considered assemblages (Krebs 1999). Zeta diversity, on the other hand, can quantify diversity for any number of sites, such as a 9 cell neighbourhood. Formally, it can be defined as "the average number of species shared by any number of sites" (Hui & Mcgeoch 2014).

To demonstrate the algorithm employed to compute zeta diversity, consider the following example (Hui & Mcgeoch 2014). Let the Venn diagram inset (Figure 2.11) show the diversity partitioning of three assemblages with species partitioned into seven disjoint sets (species), i.e. components A – G.

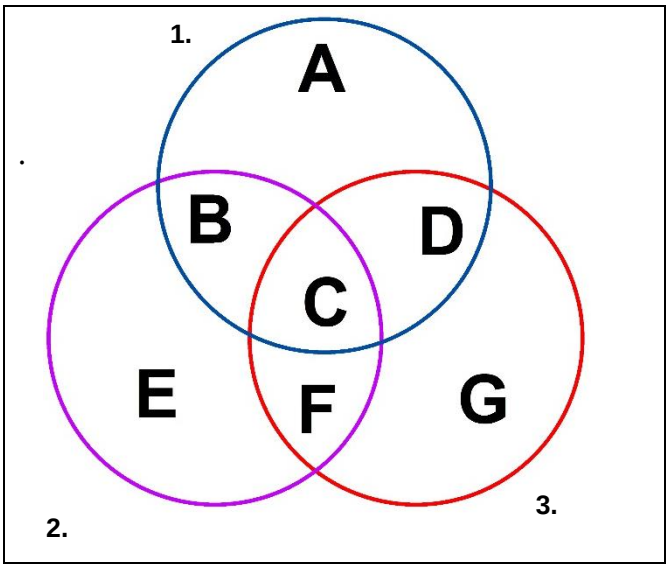


Figure 2.11: Venn diagram depicting assemblages (1-3) and species (A-G).

As an incidence-based metric, zeta diversity requires sites to be inventoried, that is, for presence and absence to be established. Standardising the data in a conventional presence-absence matrix with sites as rows and species as columns renders Table 2.4.

Table 2.4: Species presence inventory for three sites. Presence in a sampling unit is indicated by an × while absence is indicated by an *o*.

	Species						
Sites	**A**	**B**	**C**	**D**	**E**	**F**	**G**
1	×	×	×	×	o	o	o
2	o	×	×	o	×	×	o
3	o	o	×	o	×	×	×

The inventory sets for the three sites are:

Site 1 =$|ABCD|$; Site 2 =$|BCEF|$; Site 3 = $|CDFG|$

Zeta diversity can then be partitioned and computed as follows:

Let the component ζ_i be the mean number of species shared by *i* sites or assemblages, hereforth referred to as orders. For order where *i = 1*, the diversity component ζ can be computed as:

Site 1 = $|A + B + C + D|$ = 4; Site 2 = $|B + C + D + E|$ = 4; Site 3 = $|C + D + F + G|$ = 4

That is:

$$\zeta_1 = \frac{\text{Total } (|\text{Site 1}| + |\text{Site 2}| + |\text{Site 3}|)}{\text{Total Sites Considered}} = 4$$

For all purposes, zeta of the first order ($i = 1$) denotes the the mean number of species per site (species richness) or sampling unit and is the empirical equivalent of Whittaker's mean alpha diversity parameter ($\bar{\alpha}$).

Similarly, for order where i = 2 (i.e. the second order) the diversity component $\boldsymbol{\zeta}$ can be computed as:

Sites 1 & 2 = $|B + C|$ = 2; **Sites** 1 & 3 = $|C + D|$ = 2; **Sites 2 & 3** = $|C + F|$ = 4

$$\boldsymbol{\zeta_2} = \frac{\text{Total species shared}(|\text{Site (1\&2)}| + |\text{Site (1\& 3)}| + |\text{Site (2\&3)}|}{\text{Total Sites Considered}}$$

Quantitatively, $\boldsymbol{\zeta_2}$ is thus the average number of shared species between any two sites for all possible pairwise combinations of those sites. Hence, $\boldsymbol{\zeta_2}$ implicitly relates to the 'a' parameter used in similarity indices which indicates sites with co-occurring species.

Similarly, for order where i = 3 the diversity component $\boldsymbol{\zeta}$ can be computed as:

$$\boldsymbol{\zeta_3} = \frac{\text{Total species shared}(|\text{Site } (1+2+3)|}{\text{Total Sites Considered}}$$

$$\boldsymbol{\zeta_3} = |C| = 1$$

The multi-site component $\boldsymbol{\zeta_3}$,

This is true for zeta where $i = n > 1$.

Hui & McGeoch (2014) note that all incidence-based, pairwise beta diversity metrics can be expressed exclusively with ζ_1 and ζ_2 components. Though its potential for macroecological analysis has been noted by Socolar et al. (2015), the use of utility of zeta diversity partitioning as alternate means to map and produce assemblage structure remains limited. Zeta diversity provides an exciting avenue to assess assemblage structures across multiple sites and comprises the method use to map biodiversity patterns in the present study.

2.5 MODELLING DRIVERS OF DIVERISTY

As illustrated in the previous section, a perplexing array of measures exits to compute beta diversity as turnover. Even more perplexing, are the amounts of statistical measures available to test relationships among variables driving diversity. An overview of some commonly used modelling techniques is briefly described below.

2.5.1 **Spatial data**

Spatial data, such as species distributions, comprises both attribute and locational information. The locational component of spatial data compromises inferences made from traditional statistical techniques because these were developed for aspatial data only (Anselin 1989). Statistical analysis pertaining to spatial data thus necessitates a different set of statistical techniques and modelling approaches.

A growing amount of literature has been dedicated to the development of analytical techniques explicitly capable of handling the locational component of spatial data (Dormann et al. 2007; Zietz, Zietz & Sirmans 2008; Beale et al. 2010; Tsai 2011). These techniques can be broadly grouped into two classes based on the locational component of spatial data addressed, (Anselin 1989) that is:

1. Spatial autocorrelation
2. Spatial nonstationarity

The following sections (2.5.2 -2.5.3) provide an overview of both spatial autocorrelation and spatial nonstationarity and how diversity is modelled in the presence of these phenomena.

2.5.2 **Spatial autocorrelation (SAC)**

Spatial autocorrelation (SAC) refers to phenomena where, for a given variable, observations from nearby locations are more similar in size and magnitude than distant ones (Kissling & Carl 2008; Beale et al. 2010; Dobrovolski et al. 2012). Based on Toblers law, SAC shows within-variable correlation across geographic space and comprises a first-order effect in geographical analysis: “all things are related, but near things are more related than distant ones”.

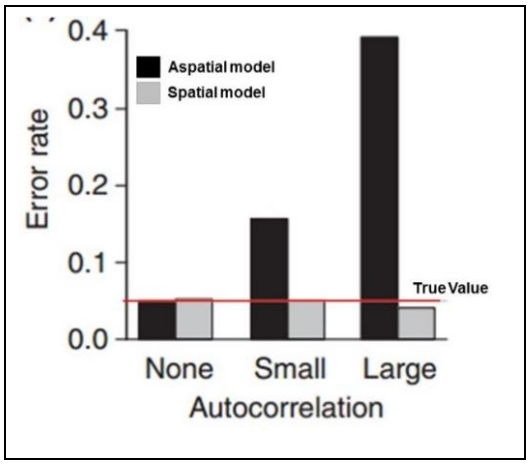
SAC can either be endogenic or exogenic depending on the processes which generate the autocorrelation (Beguería & Pueyo 2009). In the case of endogenic pattern generation, the observed spatial pattern is an inherent property of the variable of interest itself. This is referred to as inherent SAC (Fortin & Dale 2005). With respect to species diversity and species distributions, endogenous processes include dispersal capacity, biotic interactions, evolutionary processes and spatial inhibition. Endogenous processes may be entirely inherent or reflect the evolutionary response of taxa to exogenous processes.

Conversely, the identified spatial autocorrelation can be generated by exogenous processes. that is, processes independent of the variable of interest. Exogenous processes can act independently or interactively. In the latter, the interaction can either be additive when the interactions are linear or multiplicative when the interactions are nonlinear (Anselin 2002; Kissling & Carl 2008; Beale et al. 2010). Exogenous processes include environmental gradients, geomorphological processes and climatic constraints which can cause species distribution to be spatially dependent. Alternately, the observed SAC can also be caused by the statistical model used to model spatial relationships as well as the observational scale of study. Such SAC are referred to as pseudo-autocorrelation because the observed SAC are a property of the analytical processes and the scale of analysis and not necessarily the data (Beale et al. 2010).

Irrespective of the mechanism generating the spatial autocorrelation, the presence of the phenomena is undesirable and problematic for parametric statistics which assume independently distributed errors (Dormann et al. 2007). Two modelling problems arise in the presence of SAC. The first pertains to the inflation of type I error; rejecting the null hypothesis when it is true. In the presence of spatially dependent observations, classical tests

of significance can be biased and confidence intervals overestimated. The second problem arises when there are shifts in the coefficient estimates between spatial and aspatial regression models (Lennon 2000; Elith & Leathwick 2009; Saas & Gosselin 2014). Lennon (2000) notes that these shifts are proportional to the SAC of the covariate included in the model.

The aforementioned problems have implications for model interpretation and hypothesis testing of diversity-environment relationships such that in the presence of SAC, true spatial relationships may be confounded. For example, in a meta-analysis of 50 ecological publications, Dormann et al. (2007) found that more than 80% of the reviewed diversity-environment studies had modelled relationships using non-spatial methods. He further notes that of these, only 25% mentioned the implications or presence of SAC amongst the data. In a follow up to Dorman et al. (2007), Beale et al. (2010) repeated the analysis of some of those studies using spatially explicit models, concluding that error rates (type I & II) of aspatial models were up to ten times higher than those of spatial models (Figure 2.12). Follow up publications (Kissling & Carl 2008; Charlton & Fotheringham 2009; Fortin et al. 2012) have endorsed the use of spatially explicit models in the analysis of spatial data.



Source: Beale et al. (2010)

Figure 2.12: Model error rate in the presence of spatial autocorrelation. Error rates are higher in non-spatial regression models in the presence of spatial autocorrelation opposed to spatial model.

An important goal of geographical ecology is to establish the relationship between spatially structured variables. However, most geographical datasets contain SAC and traditional parametric statistics are often not suited for the analysis of such data (Dormann et al. 2007). SAC violates traditional parametric assumption of spatial independence in the error term which biases confidence levels of significance tests, thus increasing the chance of type I and type II error (Qi & Wu 1996; Beale et al. 2010; Kallimanis & Koutsias 2013).

To account for this, numerous modelling and statistical techniques have been developed to accommodate and account for spatial autocorrelation in regression models. The most commonly used spatial regression models in ecological analysis include the General Additive Models (GAM), Autoregressive models (AR) and the Generalised Least Squares (GLS) methods (*see* Table 2.5).

Table 2.5: Regression models commonly employed in the ecological investigation of species diversity which explicitly considers SAC during modelling

Technique	Study
Autocovariate regression (ACR)	(Fewster & Buckland 2001)
Bayesian Hierarchical modelling	(Guisan & Thuiller 2005)
Generalised Additive Models (GAM's)	(Anselin 2002; Beale et al. 2010)
Autoregressive Models (AR)	(Beguería & Pueyo 2009)
Generalised Least Squares (GLS)	(Chejanovski & Wiens 2014)
Spatial Eigen Vector Mapping (SEV)	(Murphy et al. 2015)

2.5.3 **Spatial nonstationarity**

Global regression models assume that the relationships being modelled are the same everywhere within the study area. That is, they assume spatial homogeneity (Foody 2004). However, species-environment relationships are seldom homogenous, varying across space and time (i.e. second order effect) (Fotheringham, Brundson & Charlton 2002). Second-order effects are caused by variation in the mean value of a process over the study area and is referred to as spatial nonstationarity. Empirically, nonstationarity is a form of model misspecification and results from intrinsic local differences in pattern-process relationships across geographic space and omitted explanatory variables in the regression model. Nonstationarity, thus, violates a primary assumption of global models. In the presence of nonstationary relationships, parameter estimates from global models, such as standard error scores, can be inflated (Foody 2004). For this reason, local regression models are consulted, in particular, the Geographically Weighted Regression (GWR) (Griffith 2012).

Section 2.5.4 provides an overview of the GWR and its application in species distribution modelling.

2.5.4 **Modelling spatial relationships using Geographically Weighted Regression**

Geographically Weighted Regression (GWR) is a local spatial regression method developed for the explicit exploration of non-stationary spatial relationships between variables across geographic space (Charlton & Fotheringham 2009). As a local regression model, GWR supplies a local regression equation and associated parameters for every observation in the study area. It differs, however, from other local models (i.e. LOWLESS) in not looking for local variation in 'data' space but geographic space by moving a weighted window over the data across the entire study area, estimating one set of coefficient values at every chosen 'fit' point (Dormann et al. 2007; Tsai 2011). The estimation of the GWR is given by Equation 2.5:

$$Y = \beta_0(u,v) + \beta_1(u,v)x_1 + \ldots + \beta_u(u,v)x_n + \varepsilon$$

Equation 2.5: Geographically weighted regression equation.

Where β_0 is the regression intercept, $\beta_1 - \beta_u$ are the coefficients intercepts of the independent variables, $x_1 - x_n$ are the independent variables and *(u, v)* are the spatial coordinates of each location and ε is the error term. For the GWR nearer observations have a greater weight than distant ones in calibrating the local set of regression parameters. The relationship about each point *i* is measured using a weighted least squares approach and is given by the expression (Equation 2.6):

$$\widehat{\beta_\iota} = (X^T W(u_i, v_i) X)^{-1} X^T W(u_i, v_i) y$$

Equation 2.6: Weighted least squares estimation.

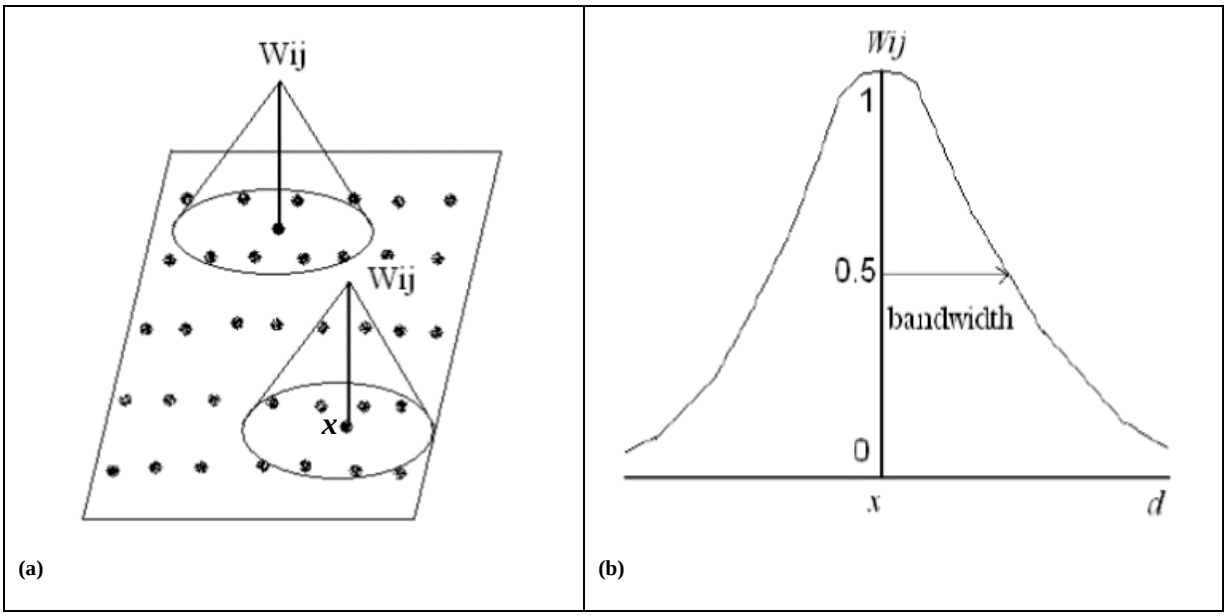
Where X^T is the matrix of independent variables, *y* is the dependent variable, *W(u, v)* is the weighting matrix of each observation whose diagonal represents the spatial weighting of observation about point *i* (Equation 2.7):

$$W(x_i) = \begin{pmatrix} w_{ij} & 0 & 0 \\ 0 & w_{i2} & \\ & & w_{i3} \end{pmatrix}$$

Equation 2.7: Standard form of spatial weights matrix of the GWR.

For GWR the weighted estimator, spatial kernel and bandwidth together determine the local model (Fotheringham, Brundson & Charlton 2002). The spatial kernel determines how the weighted estimator is applied. Several kernels can be distinguished, each of which apply a different weighting scheme during calibration (how nearby values influence each other). Kernels can be continuous or discontinuous. Continuous kernels consider the influence of all

observations on each other while discontinuous kernels consider only the influence of *n* neighbours or observations within a distance *d*, where *d* is determined by the bandwidth parameter (Figure 2.13)



Source: Fotheringham, Charlton & Brundson (2002)

Figure 2.13: Spatial kernels in GWR. The region about regression point x (a) describes the weight of influence (wij) of points j on point x. (b) The bandwidth d sets the distance about point x where weight j is influential.

The bandwidth determines the rate at which weights decrease with distance (Charlton & Fotheringham 2009). Bandwidths are either fixed or adaptive. For a fixed bandwidth, the distance remains constant. In an adaptive bandwidth, the distance can change over the study area depending on some specified optimization parameter (Cross Validation, Akaike Information Criterion (AIC)). Nthiwa (2011) notes that fixed bandwidths are best suited for regular sampling configurations and adaptive bandwidths are best suited when sampling configurations are irregular.

2.5.4.1 The use of GWR in species-diversity analysis

GWR is being increasingly used in broadscale macroecological analysis and for species distribution modelling (Foody 2004; Dormann et al. 2007; Centre-ville et al. 2012). Foody (2004) used the GWR to relate avian richness to climatic variables across a range of spatial scales. Similarly, Ortiz-Yusty, Páez & Zapata (2013) used the GWR to examine the relationship between amphibian richness and precipitation across Colombia. Other notable studies include those of Propastin, Kappas & Erasmi (2008) and Czarnota, Wheeler & Jennings (2015).

For beta diversity, the use of GWR has been limited and to the knowledge of the author, no study has used GWR to examine patterns of beta diversity. Searching the keywords “GWR”

and “beta diversity” through the Google Scholar online academic database delivered no results. Lack of usage of GWR to study differentiation diversity is interesting because the processes that shape alpha diversity are likely to shape beta diversity (Hui & McGeoch 2014), and the GWR may provide insight into how beta diversity-environment relationships are locally structured and spatially variable. Using the GWR to model beta diversity relationships thus provides a window of opportunity to explore how local spatial models can detect and relate changes in species compositions to environmental variables across geographic gradients.

2.5.4.2 Advantages of GWR with respect to this study

GWR presents several advantages for the study of diversity-environment-scale relationships (Charlton & Fotheringham 2009). First, GWR can be used as a graphical tool for data exploration including the mapping of parameters (coefficients). This can be useful for the identification of missing explanatory variables in global models. Second, GWR can be used to explore how relationships vary across spatial scale. Third, though not explicitly developed for SAC, GWR has shown to sufficiently reduce SAC for a given dataset (Sheehan, Strager & Welsh 2013) thereby reducing the chance of type I error.

2.5.4.3 Limitations of the GWR

GWR suffers from several analytical limitations. First, GWR has shown to be more sensitive to local collinearity than global models. Second, GWR does not explicitly consider the influence of SAC and thus parameter estimates (coefficients) may be prone to parameter bias. Third, GWR is computationally intensive and sensitive to large spatial datasets. Lastly, GWR does not allow for extrapolation beyond the study region and is thus limited in its predictive power. Jetz, Rahbek & Lichstein (2005) note that this limitation is perhaps why the incorporation of GWR into ecological literature has remained relatively slow. Limitations of the GWR exclusive to the study are also discussed in later sections.

2.6 MODELLING SPECIES TURNOVER

Legendre et al. (2005) proposed two approaches for analysing and modelling patterns of beta diversity: the raw data approach and the distance approach. In the raw data approach, environmental variables and geographical components of beta diversity are partitioned through regression analysis. In the distance approach, dissimilarities in biological diversities are modelled as a function of the environmental or geographical distance between sites using either matrix correlation or matrix regression.

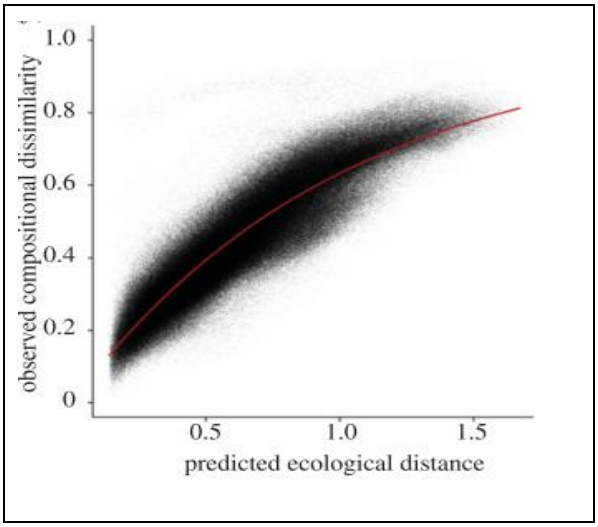
2.6.1 **Correlation approach**

In the correlation approach, the relationship between beta diversity and the environment is established through regression and correlation analysis. Commonly used regression methods include the Partial Least Squared regression and traditional linear models. For example, Baselga et al. (2012) used Partial Least Squares regression to relate species turnover in amphibians to present and paleo climatic conditions. Similarly, Zhang et al. (2014) used simple linear regression to relate the species turnover of steppe grasses to annual precipitation in Mongolia, while Veech & Crist (2007) used multiple linear regression to assess the influence of habitat heterogeneity on the species turnover of birds.

Ferrier et al. (2007) notes that although traditional regression models provide the advantage of being simple, they produce less precise model fit parameters. For this reason, ordination and partitioning techniques are often preferred to traditional regression analysis. Ordination techniques include Canonical Correspondence Analysis (CCA) and Detrended Correspondence Analysis (DCA) (Dray et al. 2012).

2.6.2 **Distance approach**

In the distance approach (Figure 2.14), biotic dissimilarity is modelled as a function of the environmental or geographical distance between sites using matrix regression. First, the species and environmental data are standardised into matrix format on a 0-1 scale where after the matrices are regressed using linear or nonlinear regression methods. The distance approach is frequently used for the hypothesis testing of spatial and niche based processes in the structuring of species turnover (Getis 1999; Wilson & Meurk 2011; Saito et al. 2015). Gotelli & Colwell (2011) note that owing to its complexity, matrix regression is recommended only if sufficient computational power is available or if the analysis is conducted for small extents with fine grain or large extents with large grains.



Source: Fitzpatrick et al. (2013)

Figure 2.14: Distance approach to modelling biotic dissimilarity.

Ferrier et al. (2007) introduced the "Generalised Dissimilarity Modelling" (GDM) approach ,which is an extension of the distance approach, but accounts for spatial dependence and non-stationarity amongst the data. Valdujo, Carnaval & Graham (2013) recently used this method to determine the environmental correlates of anuran diversity across central Brazil. Fitzpatrick et al. (2013) also used GDM to determine the species turnover of European and Australian butterflies.

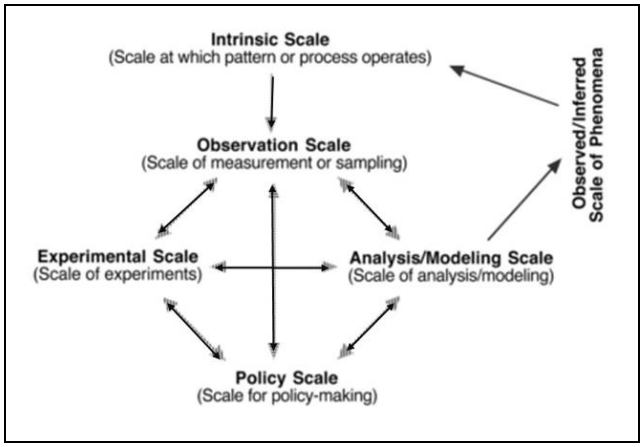
2.7 SPATIAL SCALE

2.7.1 **Description**

Beta diversity, in its broadest sense, describes variation in species identities from one site to another (Jurasinski et al. 2009). It is fundamental to both community ecology and underpins applied conservation ecology (Chave 2013). One of the most confounding factors for quantifying the patterns and processes driving diversity is spatial scale (Lustig et al. 2015).

Ecologists distinguish between several types of scales (Figure 2.15) of which the most prominent are the intrinsic (characteristic), observational experimental, modelling and analysis scale (Sale 1998; Wu & Li 2006). The intrinsic scale, or characteristic scale, refers to the scale at which an ecological pattern-processes actually operates (Wu & Li 2006). The observational scale of study refers to the scale at which sampling occurs or measurements are taken. Wu & Li (2006) suggest that the intrinsic scale is a consequence of both the interaction between the observer and the inherent scale of the phenomena and thus the intrinsic and observational scale of study are inherently related. This notion is endorsed by Chave (2013),

who postulates that "only when scales of observation or analysis are properly selected can the true characteristic scale of the phenomena of interest be detected".

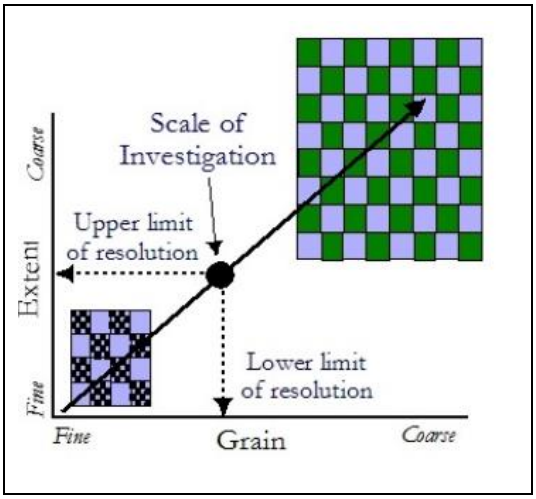


Source: Wu & Li (2006)

Figure 2.15: Kinds of scale.

The analytical and experimental scales comprise the scales at which hypothesis testing and spatial modelling occurs. Inferences and deductions about the intrinsic scale are made based on observations from the experimental and analysis scales. The analytical and experimental scales jointly determine the policy scale; the scale at which policy-making such as conservation strategies, are drawn up and implemented.

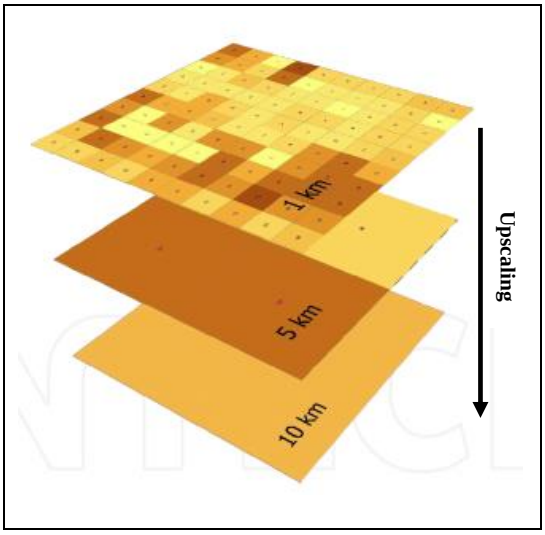
The concept of scale itself is composed of several components namely, grain, extent, lag and cartographic ratio (Levin 1992). In community ecology, grain and extent are considered the most prominent, where grain refers to either the spatial resolution of image used in the study or the size of the individual units of observation (i.e. quadrant-sample). Extent, on the other hand, refers to either the outer most bounds of a dataset or the overall area encompassed by a study (Nekola & White 1999; Loreau 2000; Rahbek 2005). Grain and extent define the lower and upper limits of resolution in the data (Figure 2.16). This is because spatial patterns cannot be detected at finer or coarser scales than the grain and extent of the data, respectively. Thus any inferences about scale-dependency in a system are constrained by the extent and grain of investigation (O'Neill et al. 1991).



Source: McGarigal (2000)

Figure 2.16: The relationship between grain and extent.

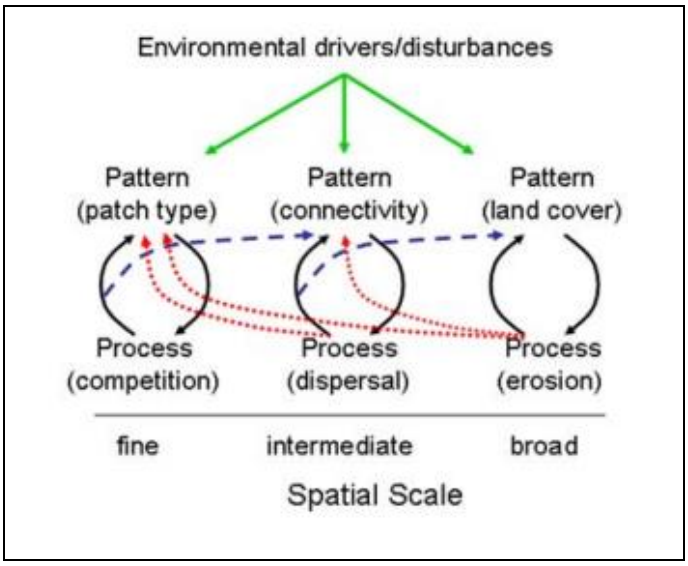
In most scale-multiplicity studies only grain size is altered while extent is held constant (Gibson & Ostrom 1998; Hui 2009). In ecology, this process is usually referred to as scaling which can be defined as "the extrapolation or empirical translation of information from one scale to another" (Francis & Klopatek 2000, Fisher & Mustard 2007). Based on the directionality of the scaling mechanism, two types of scaling operations can be distinguished: upscaling (bottom-up) and downscaling (top-down) (Turner 1989, Wu & Li 2006). Upscaling refers to the empirical transfer of information from finer to coarser scales (Ewers & Didham 2008). In contrast, downscaling consists of decomposing information from broader to finer scales (Gibson & Ostrom 1998). Upscaling (Figure 2.17) is preferred to downscaling because information can be preserved during upscaling but not during downscaling where a net loss of information occurs and inferences about detail may not be true. Since ecological phenomena exhibit different patterns at different spatial scales, scaling is a useful technique for relating and inferring processes and patterns across multiple spatial scales (Wu & Li 2006).



Source: McGarigal (2000)

Figure 2.17: Upscaling translates information from finer to coarser scales.

Much research has been done to try and establish how scale influences not only the visual perception of ecological patterns but also the metrics associated with describing these (Fisher & Mustard 2007). Three examples of scale problems in ecology include: (1) the identification of scales of patterns; (2) the effect of changing scale on patterns and; (3) the effect of changing scale on pattern-process relationships (Figure 2.18) (Gitelson et al. 2012; Chave 2013; Wang 2013).

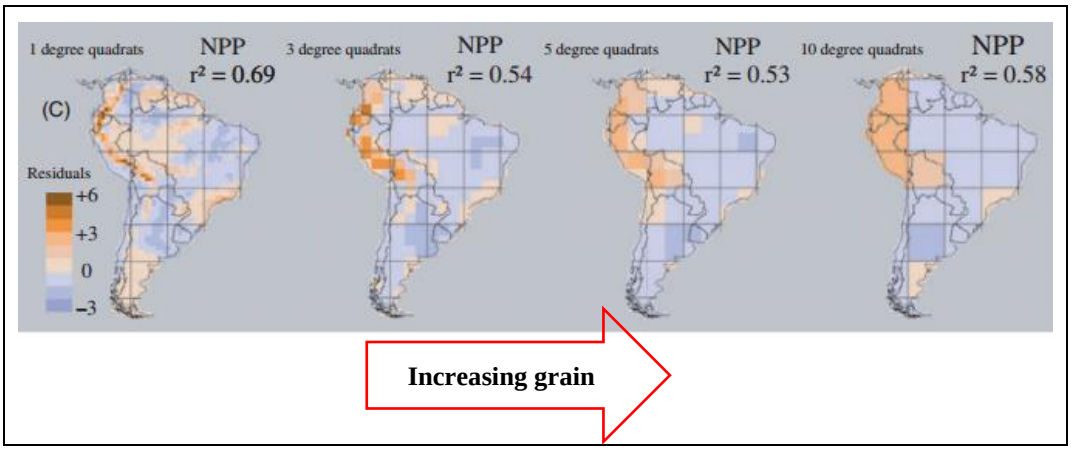


Source: McGarigal (2000)

Figure 2.18: Change in importance of exogenous and endogenous processes with changing spatial scale. The red arrows indicate how broad scale process can influence patterns at a fine scale while the blue arrows indicate how these fine scale patterns interact to produce broad scale patterns.

For example, at a fine scale, alpha diversity (α) has shown to increase rapidly with increasing extent due to high variation in stochastic species occupancy patterns among sampling units and the deterministic variation in species responses to habitat heterogeneity. At an intermediate scale, alpha diversity displays less growth while at broader scales growth again

is accelerated due to geological barriers and species having distinct evolutionary histories (Colwell 1988; Barton et al. 2013). The spatial dynamics, observed patterns and statistical relations between pattern-process in an ecological system studied at a given may thus not be actual but an artefact of its scale of study (Wiens 1989). For example, studying the relationship between South American avian richness and ecosystem productivity across multiple spatial scales, Rahbek (2005) found that patterns of avian richness and the relationship these posed with productivity were highly scale dependent, shifting with each alteration of grain (Figure 2.19). Other such observations include those of Rodríguez, Alfonso & Hawkins (2005) and Astorga et al. (2014).



Source: Rahbek (2005)

Figure 2.19: Species diversity patterns and ecosystem productivity relationships are scale dependent.

Scale and scaling relations are thus salient in understanding any aspect of ecological and geographical systems since the central premise of both disciplines comprises the identification and quantification of spatial patterns and the mechanisms shaping these (Levin 1992; Sale 1998; Hawkins et al. 2003; Wiens & Donoghue 2004).

Inherent to the issue of scale is the Modifiable Areal Unit Problem (MAUP), a form of statistical bias which occurs when area based data are aggregated (Marceau & Hay 1999). The MAUP comprises both a scale and zone effect (Francis & Klopatek 2000). This means that though the underlying spatial information remains the same, the way in which it is aggregated (i.e. zone or scale) directly affects the way it is interpreted both visually and statistically. For example, as grain is increased within sample heterogeneity and variance is decreased, resulting in a progressive smoothing effect at coarser grains. This smoothing effect is the underlying cause of both the scale problem in the MAUP and aggregation bias in ecological studies often known as the “ecological fallacy” phenomena.

2.7.2 **Scaling and diversity: A conceptual framework**

Cushman, McGarigal & Neel (2008) note that in order to identify the characteristic or intrinsic scale of a phenomena, scaling techniques need to be employed. Scaling can reveal or reveal multiple agents operating at different scales. Several scholars (Wiens 1989; Wu 2004; Ewers & Didham 2008) have suggested that all scaling relationships, regardless of the variable employed, will exhibit one or a combination of three types of behavioural patterns when grain size is progressively aggregated to larger sizes, these are:

1. Predictable linear, power or log responses;
2. Staircase like responses; and
3. Erratic responses exhibiting unpredictable and inconsistent scaling relations

Type one behaviour suggests that most modelled phenomena will change in a linear, power or log like manner as a function of grain. Type two behaviour suggests that a change in scale will result in a stepwise change in pattern or process (i.e. scale breaks). This type of behaviour is cemented in hierarchy theory which posits that "ecological pattern-processes occur at distinct scales, due to the nonlinear interaction of biotic and abiotic components, resulting in distinct scales of spatial patterning" (Peters, Bestelmeyer & Turner 2007). Simply stated, it suggests that discrete changes in grain or extent, will result in discrete changes in spatial patterns, process and pattern-process interrelations (Chave 2013). Type three behaviour suggests that no consistent or discernible scaling pattern exists for the observed phenomena across spatial grain.

Furthermore, Rahbek (2005) notes that there is no universally correct scale of analysis. Rather, an optimal scale of analysis for a particular pattern-process can be identified by systematically varying the scale of analysis from finer to coarser scales. Blackburn & Gaston (2002) support this notion postulating that in order to understand ecological pattern-process relations, multi-scale studies need to be conducted if the agents structuring diversity are to be disentangled.

Following a meta-analysis of scale-diversity literature, Rahbek (2005) proposed a conceptual framework, comprising six elements or recommendations, for conducting multiscale studies of ecological pattern-process relationships:

1. Resample or aggregate data across multiple grain sizes to explore patterns and causal interrelationships between variables;
2. Diversity studies must be conducted across large spatial extents with distinct but comparable geographical gradients;
3. Scale studies must observe both pattern and pattern-process relationships because the two are not mutually exclusive;
4. To establish statistical relationships, traditional parametric statistics and models should be abandoned because spatially structured processes are spatially autocorrelated which violates the assumption of spatial independence of these statistics thereby increasing the chance of type I error ultimately leading to bias or erroneous conclusions across scale;
5. With respect to point four, spatial regression models should be used as an alternative; and
6. With respect to point five, both global and local approaches are recommended, the results of which must be presented as function of spatial scale;

This approach was used to model diversity across multiple scales in the present study.

2.7.3 **The spatial scaling of species diversity: Alpha and beta diversity**

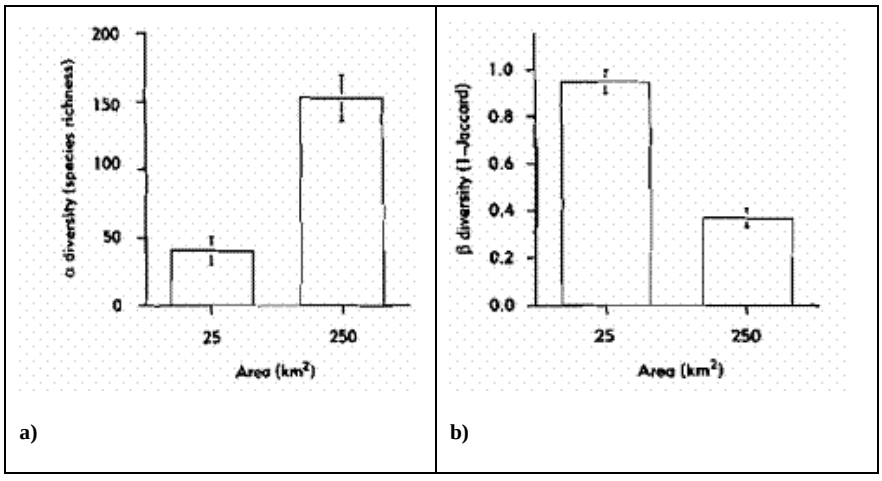
Few studies have focused solely on the scaling relations of beta diversity (Barton et al. 2013). As such, no general framework exists for the scaling of species and the absolute scales at which alpha and gamma diversity should be measured continue to remain elusive. One of the reasons for this is that species perceive and respond to the world at different scales. Selecting an appropriate scale of study thus remains problematic. Currently, the selection of spatial grain and extent of sampling are strongly influenced by both the biology and geography of the taxon of interest, particularly their presumed dispersal capacity (Sale 1989; Barton et al. 2013). For example, soil faunal communities are often quantified in sampling units of a few square centimetre while mammal communities are quantified in units of a few kilometres (Woodcock & Strahler 1987; Sale 1989).

Variation in diversity or species composition at local, regional and global scales can be attributed to different processes operating at, and resulting in the manifestation of, different patterns at different scales. Fine scale studies reveal greater detail about biological mechanisms underlying patterns of phenomena of interest, whilst generalizations are more

likely to occur broader scales (Wu & Li 2006). As the reader may recall, spatial variation in beta diversity is usually conceptualised using two approaches – first, the partitioning of species diversity into alpha and gamma components as made by Whittaker (1960) and Lande (1996). Second, distance decay of similarity measures which regress pairwise measures of sample similarity as a function of distance and parameterize slope gradient to indicate relative change in compositional similarity through geographic space (Barton et al. 2013).

Whittaker's initial beta diversity formula (βw) is insensitive to species richness but not to individual numbers (Wiens 1989; Vellend 2001). As such, it is regarded as a less appropriate measure for conducting studies across scales of turnover (Barton et al. 2013). Barton et al. (2013) suggest that studies scaling beta diversity as turnover should instead employ normalised measures of differentiation, such as similarity indices, since they account for variation in the number of communities. Such measures also have the advantage of supplying an average among-sample dissimilarity at a specified scale and are thus useful to consider when comparing different taxa or regions with varying levels of species richness (Barton et al. 2013). Studies by several scholars (Barton et al. 2013; Buckley & Jetz 2008; Whittaker, Waillis & Field 2001) have revealed that alpha diversity increases characteristically with area sampled as new species are more frequently encountered.

Currently, there is no overall consensus as to how beta diversity is affected by spatial scale and results between have not been equivocal with some studies showing beta diversity to increase with increasing grain (Navarro-Sigüenza & Rodríguez 2014) while others have shown a decline in beta diversity with increasing grain (Keil et al. 2012) (Figure 2.20).



Source: Magurran (2004)

Figure 2.20: Scale dependency of (a) alpha and (b) beta diversity across spatial grain.

Scaling beta diversity, Keil et al. (2012) suggested that patterns and scaling relations of diversity are likely to be dictated by the species-area relationship and behave in three ways.

First, biotic dissimilarity between adjacent cells is likely to increase as function of grain owing to the species-area relationship. This is because as grain size is increased, the area it encompasses also increases whereby more rare species are likely to be included in each sampling unit resulting in greater compositional dissimilarity between adjacent sampling units. Second, as grain is increased, within sample heterogeneity and variance are decreased resulting in higher beta diversity due to smoothing of data. Lastly, beta diversity is likely to be higher at coarser grain because as grain size is increased, the distance between sampling units also expands. Consequently, this leads to an increase in environmental and spatial distance between adjacent units resulting in higher beta diversity. Beta diversity is thus expected to be higher at coarser grains for species with low dispersal capacity and environmental tolerances such as amphibians.

Blackburn & Gaston (2002) note that ultimately the scaling of beta diversity will depend on the measure used to quantify beta diversity, the biogeography and evolutionary history of organism and phenomena of interest, the scale of analysis and the statistical parameters used to quantify and relate pattern to process.

2.7.4 **Contribution of GIS and Remote Sensing to scaling studies**

In the past, ecological analyses were largely confined to small spatial grains and extents. Reasons for this included a lack of sophisticated tools and computational power to conduct studies at broader spatial extents as well as the failure to acknowledge the influence of scale on observed spatial patterns-processes. With the rise of the Geographic Information Sciences (GIS) and advancement in Remote Sensing technologies in the 1960s-1970s as well as the acknowledgement of the scale problem within these disciplines (Marceau & Hay 1999), ecologists were able to conduct studies at much broader spatial extents.

In a review of spatial scale in ecology and geography, Marceau & Hay (1999) noted the contribution of the GIS and spatial statistics developed for spatial analysis, such as interpolation and aggregation techniques, for being the essential tools needed to store, manipulate and model spatial patterns and processes across a range of spatial scales. For example, remotely sensed climate data over broad spatial extents and GIS interpolation techniques have jointly allowed for the derivation of climatic data for previously inaccessible areas.

Furthermore, remotely sensed images and derived spectral-environment proxies, such as the Normalised Difference Vegetation Index, have allowed for studies pertaining to changes in

landscape patterns to be conducted across a range of spatial grains. Francis & Klopatek (2000) note that the advances in remote sensing and GIS technology have allowed for "ecologists to scale up, not only in order to understand the complex and dynamic interactions between land and atmosphere, but also to try and convey information about fine scale ecological patterns and processes to broad scale applications".

2.8 CONCLUSION

In conclusion and summation of the literature review, several key deductions can be made. First, amphibians are a taxa of prime conservation concern with global populations declining at an unprecedented rate. Second, numerous factors influence diversity with many competing environmental hypotheses formulated to explain diversity gradients. Third, there is no single correct way to measure biological diversity or model diversity-environment relationships. Rather, the measures employed or methods followed must reflect the purpose of the study as well as the research aims and objects. Lastly, spatial scale is integral in understanding spatial patterns, spatial processes and pattern-process interrelations and thus needs to be explored if the true mechanisms driving diversity gradients are to be adequately understood.

CHAPTER 3: DRIVERS OF REGIONAL RICHNESS AND TURNOVER

This chapter comprises the first analytical component of the research document. Specifically, it examines the relationship between species diversity and the environment. Chronologically, it comprises a description of the data, the research methods followed, the results obtained and a discussion linking the results to existing literature.

3.1 DATA COLLECTION: OVERVIEW

Two types of data were employed in this study: species distribution data and environmental data. The species data were in the form of vector shapefiles and the environmental data were in continuous raster format.

The study was conducted using several software packages. Statistical analysis was conducted using IBM SPPS 23.0 and XLSTAT 4.0. Regression analysis was conducted using the Spatial Analysis for Macroecology (SAM) 4.0 and ArcGIS 10.3.1 (ESRI 2015) suites. All maps were generated using ArcMap 10.3.1 while statistical figures in the analytical components (graphs, tables) were created using the Microsoft Excel 10.1 statistical package.

Projections and geographical coordinate systems constitute salient components of spatial analysis as they determine the location and manner in which objects are presented on the earth's surface. Different projections preserve different properties of spatial data and thus differ in their spatial and empirical consequence. In the present study, all data were projected using an Africa Albers Equal Area Conic projection. This co-ordinate system was selected because: it has been tailored to preserve the area of the study extent, reduces sampling bias as data are extrapolated to coarser grains allowing for hierarchal sampling and is in keeping with established ecological literature.

3.1.1 Species data collection

Global distribution maps of 6157 extant species of amphibians were obtained from the IUCNs 'Red List' species data archives (<http://www.iucnredlist.org/technical-documents/spatial-data>) (Sodhi et al. 2008). The data were in the form of distribution polygons, each representing the extent of occurrence (EOC) for a species (Figure 3.1). The data were compiled by the IUCN Global Amphibian Assessment (GAA) unit over 13 years (2000 – 2013) and were validated by renowned ecologist Michael Hoffmann.

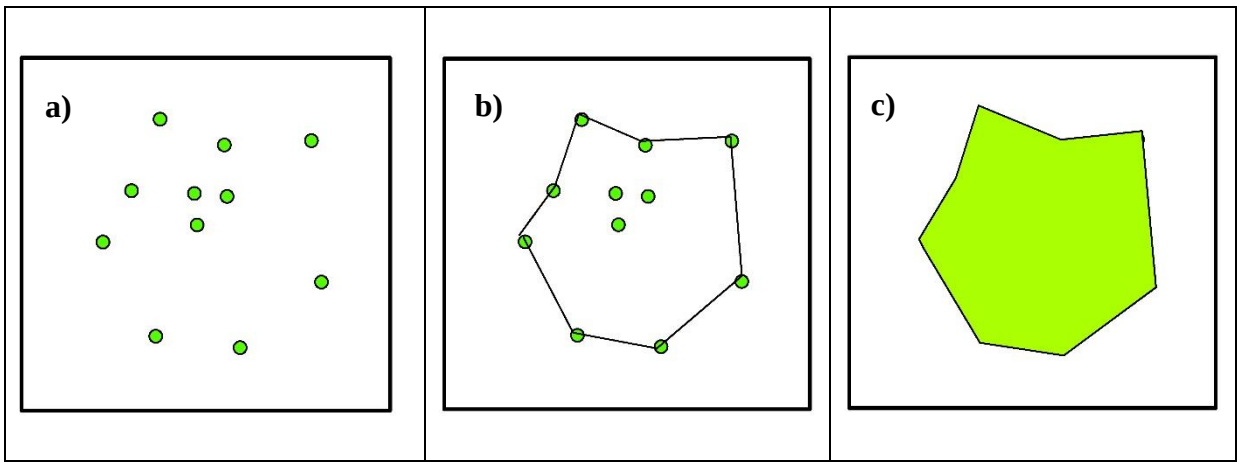


Figure 3.1 Amphibian distribution polygons as compiled per IUCN methodology: (a) nodes represent the spatial location of known, inferred or projected sites of species occurrence. The boundary to the extent of occurrence (b) is estimated by a plotted convex hull which is used to derive (c) the spatial distribution polygon for each species.

In an initial pre-processing step, the data were assembled into ESRI's ArcMap where it was checked for logical inconsistencies and geometric errors using the software's built-in "check geometry" tool. Checking for sampling and digitizing artefacts constitute key steps in spatial data management and were done to avoid inducing sampling bias into analysis. Several digitising errors, comprising sliver polygons and overshoots, were noted in the dataset and were subsequently repaired. The repaired dataset was then clipped to the study extent where anuran data were extracted leaving 726 species with 1158 records for further analysis.

3.1.2 **Species data: Scale and sample size**

Species presence and absence was then established using a grid system comprising 5943 contiguous equal area quadrats roughly equivalent to 0.5º × 0.5º (56 × 56 km) near the equator using an equal area projection. Equal area grid systems comprise the standard sampling scheme in broad scale macro-ecological analysis and are preferred to hexagonal lattices because their geometry are less sensitive to geographical transformations, avoid area biases and allow for hierarchical structuring (Whittaker, Willis & Field 2001).

The species dataset was then assembled in ArcGIS where presence was established by superimposing the grid system over the species dataset. A species was considered present in a quadrat if any portion of its geographic range polygon intersected a quadrat at any locality (McKnight et al. 2007; Melo, Rangel & Diniz-Filho 2009). A presence/absence matrix of the recorded occurrences was then compiled to allow for the computing of diversity metrics using functions written in the *R* 3.2.3 statistical environment (Appendix A). The matrix, with rows as sites and columns as species, was set in binary format with presence encoded as 1 and absence as 0.

3.1.3 **Quantifying diversity**

Zeta diversity (ζ) was then used to quantify patterns of diversity. Zeta diversity is a recently formulated (Hui & Mcgeoch 2014) incidence-based metric that captures all diversity components produced by assemblage partitioning. It measures the average number of species shared by any number of sites where the component ζ_i is the mean number of species shared by *i* sites referred to as orders. The first order (ζ_1) denotes the mean number of species per site and is the empirical equivalent of Whittaker's mean alpha parameter ($\bar{\alpha}$). The second order zeta component (ζ_2), is the mean number of species shared by two sites and is the ***a*** parameter used in compositional similarity indices of diversity. Zeta diversity was selected because unlike traditional pairwise measures, the metric can capture assemblage partitioning across a range of spatial scales for any number of sites.

In the present study, alpha diversity (α) was considered as a synonym for species richness and defined as the mean number of species in a sampling quadrat. The first order zeta component was used to quantify alpha diversity (ζ_1). Beta diversity, as turnover, was defined as the average dissimilarity between a focal cell and an adjacent cell. To measure beta diversity as turnover, several of the 24 beta diversity indices reviewed by Koleff, Gaston & Lennon (2003) were considered. Ultimately, the Sorensen dissimilarity index was selected because of its simplicity, high rating in ability to capture biotic heterogeneity and wide use in beta diversity studies (Legendre 2008; Dobrovolski et al. 2012; Réjou-méchain & Hardy 2012). In addition, Barton et al. (2013) note that the normalised Sorensen is ideal to observe patterns of beta diversity across spatial grain because of its insensitivity to the 'true' number of communities.

The Sorensen index (β_{sor}) uses three mismatching components to describe a pairwise change in compositional similarity: ***a*** = the number of co-occurring species between two quadrats, ***b*** = the number of species exclusive to the focal quadrat and ***c*** = the number of species exclusive to the adjacent quadrat (Equation 3.1) (Baselga, Gomez-Rodriguez & Lobo 2012).

$$\beta_{\text{sor}} = 1 - \frac{2a}{2a + b + c}$$

Equation 3.1: The Sorensen dissimilarity index.

In terms of partitioned zeta components (Equation 3.2), the parameters for the Sorensen index can be interpreted as $\mathbf{a} = \zeta_2$, $\mathbf{b} = \mathbf{c} = \zeta_1 - \zeta_2$ and the index becomes:

$$\zeta_{sor} = \frac{2\zeta_2}{2\zeta_2 + (\zeta_1 - \zeta_2) + (\zeta_1 - \zeta_2)}$$

$$\zeta_{sor} = \frac{2\zeta_2}{2\zeta_2 + 2\zeta_1 - 2\zeta_2}$$

$$\zeta_{sor} = \frac{2\zeta_2}{2(\zeta_2 + \zeta_1 - \zeta_2)}$$

$$\zeta_{sor} = 1 - \frac{\zeta_2}{\zeta_1}$$

Equation 3.2: The Sorensen dissimilarity index expressed in terms of zeta diversity components.

A primary limitation of pairwise indices is that they do not preserve information on the identity of species shared across more than two sites, consequently ignoring possible patterns of co-occurrence (Baselga 2013; Ricotta & Pavoine 2015). For this reason, both pairwise and multisite (i = 3) Sorensen indices were computed. Spearman rank coefficients (ρ) were then used to describe the correlation between the pairwise and multisite indices. When compared (Figure 3.2), the two indices were found to be highly correlated (ρ = 0.96, R^2 = 0.99). For this reason, results for pairwise Sorensen indices were considered as sufficient.

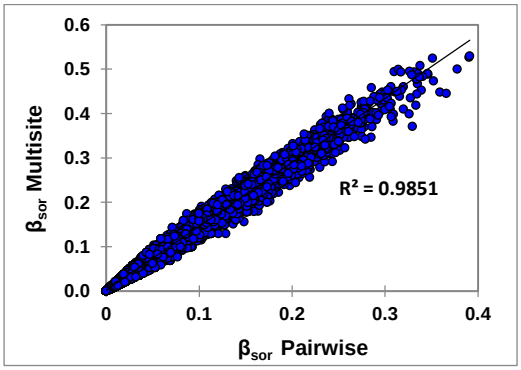


Figure 3.2: The correlation between pairwise and multisite Sorensen indices computed per zeta diversity partitioning using a spearman rank coefficient.

The Sorensen index has been criticised for being sensitive to local variations in species richness which can influence observed patterns of beta diversity (Dobrovolski et al. 2012). However, a bivariate Pearson correlation scatter gram (Figure 3.4) plotted for alpha and beta diversity metrics found the two components to be significant ($p < 0.05$) but weakly correlated (r = 0.16, r2 = 0.03). Species richness gradients were thus not expected to have a profound influence on beta diversity patterns.

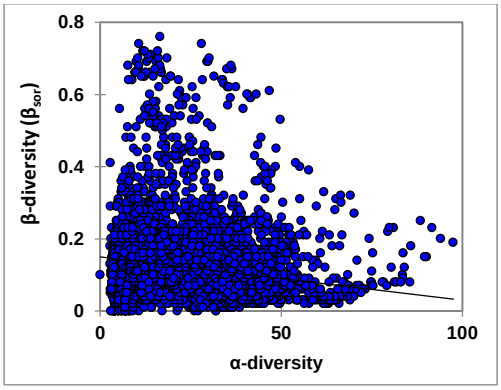


Figure 3.3: Bivariate Pearson scatter plot depicting the correlation between alpha and beta diversity.

3.1.4 **Quantifying diversity: Moving window**

Next, a moving window analysis was used to quantify diversity across extent (Figure 3.4). Specifically, for a focal cell *i*, a moving window was defined as a circle *c* with radius *r*, with its centre at *o*, the centroid of *i*. The radius *r*, was set as half the distance of the diagonal between the centroids of three contiguous sampling units: corresponding to ±80 km in Euclidean space. Only cells within distance *r* were considered as neighbours of *i*. A window thus comprised a neighbourhood of 9 cells, one focal and eight neighbours.

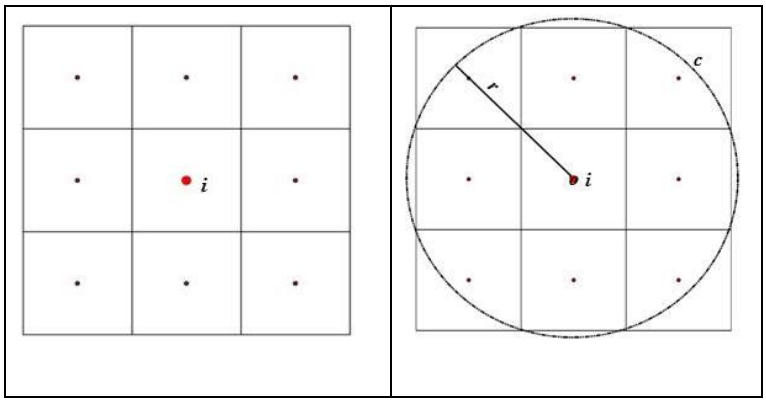


Figure 3.4: Moving window used to compute diversity across the spatial extent.

Within each neighbourhood, diversity for a focal cell was derived using a weighted average approach. That is, the diversity value assigned to a core cell was computed by using the weighted average of its eight first order neighbours (Figure 3.5). This approach for deriving diversity is similar to that of Williams, De Klerk & Crowe (1999) and is useful for the detection of anisotropic spatial patterns and deriving diversity in the absence of exhaustive sampling.

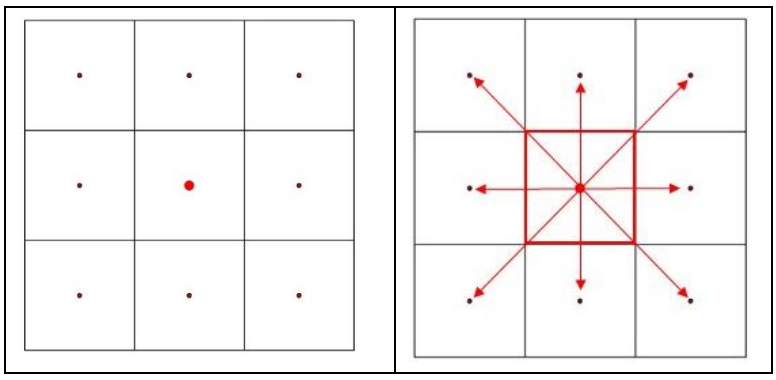


Figure 3.5: Diversity was quantified using a weighted average approach.

Window movement was unidirectional along an east west latitudinal transect (Figure 3.6). For every iteration of the window, zeta diversity was quantified through orders 1 to 3 using the *zetadiv* package as developed for *R* (<cran.r-project.org/web/packages/zetadiv/>).

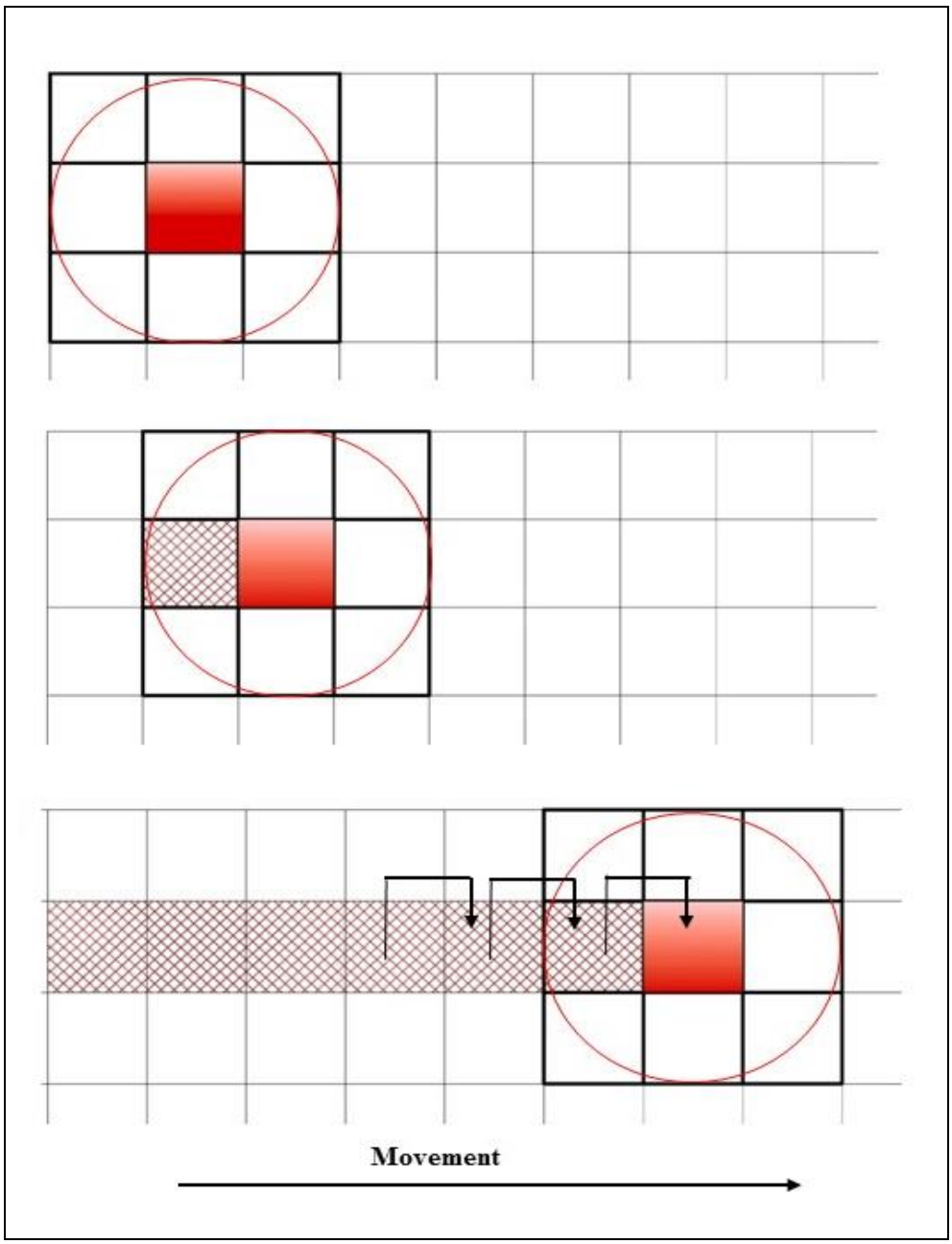


Figure 3.6: Movement of window across extent was unidirectional along latitudinal transects.

No adjustments were made for coastal cells or partially occupied neighbourhoods ($n < 8$) during spatial analysis. Consequently, values computed for such neighbourhoods may exhibit bias by either inflating or underestimating the number of species occurrences creating possible edge effects. The number of partial neighbourhoods, however, was small with respect to complete ones and was thus not excepted to have any profound impact on statistical analysis.

Once computed, maps for alpha and beta diversity were generated using ArcGIS. These were used to observe the spatial distribution of both diversity components across the study extent.

3.1.5 Environmental data description: Variable selection

Next, twenty environmental variables comprising contemporary climate, climatic variability, environmental energy, ecosystem productivity and environmental heterogeneity conditions were used to evaluate the effect of environmental processes on diversity. The selected variables are those hypothesised to govern vertebrate diversity across a range of spatial scales, realms and biogeographic regions.

For contemporary climatic conditions, mean annual precipitation (MAP, mm/y) was used as a measure of water availability and mean annual temperature (MAT, °C) as a measure of ambient energy. Climatic variability variables included temperature seasonality (TSEAS, standard deviation of mean monthly temperature) and precipitation seasonality (PSEAS, Coefficient of variation of month precipitation (CV)). Actual evapotranspiration (AET, mm/y) and potential evapotranspiration (PET, mm/y) were used as surrogates for water-energy dynamics and environmental energy. Ecosystem productivity metrics included the remotely sensed vegetation indices, net primary productivity (NPP, $g.C.m^{-2}yr^{-1}$) and the maximum Normalised Difference Vegetation Index (NDVI). Mean elevation (m) and topographic complexity, as measured by the standard deviation of each sampling quadrat, were considered as measures of physiology and habitat heterogeneity. As species are intrinsically adapted to their environments, a change in the environment is expected to result in a change in species compositions (beta diversity). For this reason, the standard deviation (SD) of each variable were also included in the analysis (= 20 variables).

The environmental datasets were acquired from several online resources. Climatic data were obtained from WorldClim (<http://www.worldclim.org/>) and represent interpolated bioclimatic surface variables (Hijmans et al. 2005). Environmental energy data are a product of the National Aeronautics and Space Administration (NASA), and are derived from the Moderate Resolution Imaging Spectrometer (MODIS), and are freely available at (<http://www.ntsg.umt.edu/project>). Productivity variables are remotely sensed proxies of ecosystem energy and plant productivity derived using spectral analysis (manipulation of spectral bands) and were obtained freely from the University of Montana's open source GIS portal (<http://www.ntsg.umt.edu/project>). Elevation data were acquired from the 'DIVA GIS' spatial data portal (<www.diva-gis.org>) and reflect interpolated digital surface terrain data as derived at 90m resolution per Shuttle Radar Topography Mission (SRTM).

All environmental datasets were acquired in a gridded format with a spatial resolution of roughly 1 kilometre square (km^2). Temporal resolution for all datasets were in excess of 10 years, which

are sufficient for inferring mean environmental conditions persisting across a region across a region. Descriptions of the environmental data are supplied in Table 3.1.

Table 3.1: Description of candidate environmental variables including unit of measurement, spatial resolution and temporal resolution.

Variable	Acronym	Unit of measurement	Description	Spatial resolution	Temporal resolution	Source
Mean Annual Precipitation	**MAP**	(mm/y)	This is the sum of all total monthly precipitation values	~1km^2	(1950 – 2000)	http://www.worldclim.org/current
Mean Annual Temperature	**MAT**	°C	Mean annual temperature of all months calculated by averaging the daily min and max of every day for a month over 12 months			
Temperature Range	**TRange**		Difference between absolute maximum temperature of the warmest month and the mean absolute lowest temperature of the coldest month			
Precipitation Seasonality	**PSEAS**	CV	This index is the ratio of the standard deviation of the monthly total precipitation to the mean monthly total precipitation (also known as the coefficient of variation) and is expressed as a percentage			
Temperature Seasonality	**TSEAS**	SD	The amount of temperature variation over a given year based on the standard deviation (variation) of monthly temperature averages			
Actual Evapotranspiration	**AET**	mm/y	The ratio between the sum of the mean monthly terrestrial evaporation and plant transpiration to the atmosphere		2000 - 2013	http://www.ntsg.umt.edu/project/mod16
Potential Evapotranspiration	**PET**	mm/y	The ratio between sum of the mean monthly terrestrial evaporation and plant transpiration to the atmosphere			
Net Primary Productivity	**NPP**	g-C m^{-2} yr^{-1}	The rate at which an ecosystem accumulates energy or biomass, excluding the energy it uses for the process of respiration			http://www.ntsg.umt.edu/project/mod17
Normalised Difference Vegetation Index (Max)	**NDVI**	Spectral ratio	Ratio of red and infrared spectral bands. Indicator of peak chlorophyll content		2001 - 2012	http://landcover.usgs.gov/green_veg.php
Mean Surface Elevation	**TOPO**	(m)	Mean surface elevation			http://srtm.csi.cgiar.org/

3.1.6 **Environmental data: Aggregation**

To avoid precision errors during analysis the environmental variables, obtained at approximately 1 km^2 spatial resolution, were scaled to precisely 1 km^2 spatial resolution using the nearest neighbour resampling algorithm (*sensu* Fjeldsa, Lambin & Ivlertens 1999). Resampling is the process of interpolating new cell values when transforming continuous raster data to a new coordinate space. The nearest neighbour algorithm is a type of resampling algorithm which retains the original cell value but within one error.

The resampled data sets were then aggregated to the size of the grid systems using the mean aggregation algorithm to derive mean values for each quadrat. To derive the standard deviation of each variable, a model was developed in ArcGIS using the model builder tool. In the model, values for aggregated SD variables were calculated as the average standard deviation of all 1 × 1 km cells within a larger half degree cell. This approach at quantifying the environment is in keeping with those of Rodríguez, Alfonso & Hawkins (2005) and Melo, Rangel & Diniz-Filho (2009).

Following aggregation, a point system was assigned to the grid system comprising a single point or centroid per quadrat across extent. These centroids were assigned Euclidean coordinates and then used to extract data values for each aggregated environmental layer. The extracted data were tabulated and converted to a .csv file format to allow for further statistical treatments.

3.2 DATA EXPLORATION AND MANIPULATION

The relationship between diversity and the environment was examined using Multiple Linear Regression (MLR) analysis. In the regressions, alpha and beta diversity were the response variables and the environmental variables were the predictors, henceforth referred to as such. However, because of the large number of predictor variables considered in the study, straightforward interpretation of simple or multiple regression models was precluded. For this reason, three preliminary data analyses were conducted in order to decide the best set and form of variables for the regression models. The preliminary analysis included an explanatory data analysis, a correlation analysis and model selection procedures. A description of the analysis is supplied in the succeeding sections.

3.2.1 **Exploration and manipulation**

In an initial explanatory data analysis step descriptive and skewness statistics were extracted for both predictor and response variables. The extracted descriptive statistics included measures of central tendency, skewness, kurtosis, Shapiro-Wilko parameters as well as frequency histograms (Table 3.2). Frequency histograms were used to observe the relative distribution of the data

while skewness and kurtosis statistics were used to measure degree of symmetry. Shapiro-Wilko tests were used to statistically test for non-normal distributions. This method checks for normality using the null hypothesis at the specified confidence ($p < 0.05$) interval. The null hypothesis here is that data are normally distributed. If the computed p-value is less than the chosen significance level, the null hypothesis is rejected and the data are not normally distributed (Graham 2003).

Traditional linear models do not make explicit assumptions about the underlying distribution of predictor variables, however, highly skewed variables can reduce model performance and bias parameter estimates by increasing residual variance. Thus, to homogenise residual variance and linearize relationships, highly skewed variables (skewness > | 1 |) were $\log_{10}$ transformed. The transformed variables included both response variables, MAP, TRange, TSEAS, NPP, NDVI and all standard deviation (SD) predictor variables. Following the transformation, the normality of the variables was increased and skewness decreased (Figure 3.7, Appendix B).

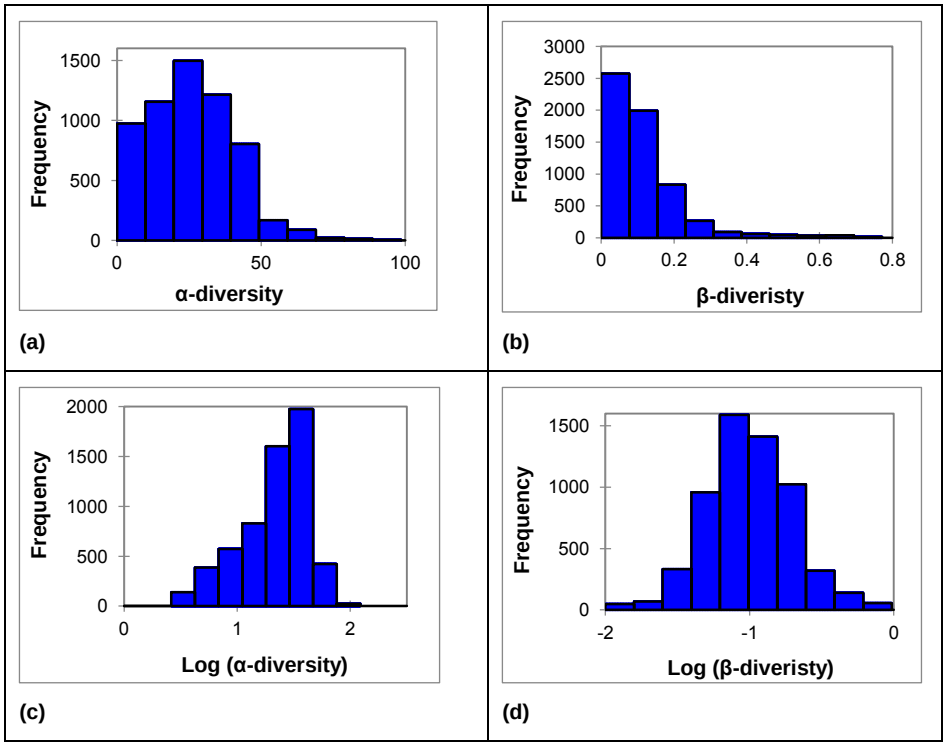


Figure 3.7: Frequency histograms of response variables, alpha and beta diversity, before (a-b) and after (c-d) $\log_{10}$ transformation.

Table 3.2: Descriptive statistics extracted for response and predictor variables.

Variable	Number. of observations	Minimum	Maximum	Median	Mean	Skewness (Pearson)	Kurtosis (Pearson)	Shapiro -Wilko
Response								
Alpha diversity	5943.00	3	97.50	26.1	26.3	1.56	0.44	P < 0.01
Beta diversity	5943.00	0	0.78	0.088	0.12	2.63	9.266	P < 0.01
Predictors								
MAP	5943.00	11.69	4189.77	973.42	1009.46	1.82	1.54	P < 0.01
MAT	5943.00	7.72	30.00	23.17	22.54	-0.65	-0.01	P < 0.01
TRange	5943.00	2.26	33.86	6.78	8.52	1.46	2.21	P < 0.01
PSEAS	5943.00	10.92	170.72	86.85	85.42	-0.08	-0.54	P < 0.01
TSEAS	5943.00	0.21	6.12	1.56	1.90	1.03	0.55	P < 0.01
AET	5943.00	0.00	146.39	63.26	64.46	0.17	-1.13	P < 0.01
PET	5943.00	0.00	312.16	235.63	234.36	-0.88	3.12	P < 0.01
NPP	5943.00	0.00	17756.40	5500.87	5523.38	0.28	-1.80	P < 0.01
NDVI	5943.00	0.00	99.91	90.27	79.04	-1.57	1.42	P < 0.01
TOPO	5943.00	-17.49	2890.48	607.34	720.32	0.76	0.44	P < 0.01
MAPSD	5943.00	0.00	1858.40	29.55	73.83	5.89	44.72	P < 0.01
MATSD	5943.00	0.00	14.26	0.50	1.45	3.08	8.58	P < 0.01
TRangeSD	5943.00	0.00	13.29	0.31	0.90	4.12	17.45	P < 0.01
TSEASSD	5943.00	0.00	1947.61	52.70	126.40	3.88	15.88	P < 0.01
PSEASSD	5943.00	0.00	83.11	1.73	5.46	3.88	15.26	P < 0.01
AETSD	5943.00	0.00	6925.07	825.72	1109.83	2.29	6.67	P < 0.01
PETSD	5943.00	0.00	15093.20	1480.87	3008.53	1.60	1.64	P < 0.01
NPPSD	5943.00	0.00	24543.00	2859.18	4754.27	1.54	1.89	P < 0.01
NDVISD	5943.00	0.00	49.80	5.21	8.15	2.53	6.52	P < 0.01
TOPOSD	5943.00	0.00	171.57	13.17	19.43	2.25	6.63	P < 0.01

3.3 MODEL SPECIFICATION

3.3.1 Correlation analysis

Following the transformations, Pearson coefficient (r) matrices were extracted for predictor variables to account for possible spurious correlations (Fortin & Dale 2005). Correlation coefficients describe the nature and degree of linear correlation between variables in a bivariate space on a ±1 scale (Legendre & Legendre 1998). A score of -1 indicates complete negative linear correlation, 0 is no correlation and +1 is complete positive linear correlation. Highly correlated variables can compromise inferences made from regression models by inducing collinearity which can bias model outputs (i.e. pseudocorrelation). In the present study, variables with $| r | > 0.7$ were considered highly correlated (Wheeler & Tiefelsdorf 2005) and systematically omitted with explicit consideration of ecological prominence to the taxon of interest (Table 3.3).

The omitted variables included AET, NPP, NDVI, TSEAS, TRangeSD, PETSD, NPPSD (Table 3.4). AET, NPP, NDVI and TRangeSD were excluded owing to high correlations with MAP. Hawkins et al. (2003) found precipitation to be paramount in determining amphibian richness due to survival and reproductive needs (laying of eggs, tadpoles). For this reason, MAP was retained.

Similarly, TOPO co-varied strongly with MAT. As solar ectotherms, anurans require sufficient amounts of ambient energy for mobility and metabolic activity (Buckley & Jetz 2007; Ortiz-Yusty, Páez & Zapata 2013). MAT was thus preferred to TOPO.

TSEAS was found to be highly correlated with TRange. Exploring determinants of amphibian richness in China, Qian et al. (2007) found that, in relation to seasonality variables, extremes in temperature variables better explained variation in amphibian diversity due to the restrictions they impose on range size (amphibians have low environmental tolerances). For this reason, TRange was preferred to TSEAS.

Lastly, PETSD was found to be highly correlated with NPPSD. Several studies (Qian et al. 2007; Calderón-Patrón et al. 2013) have found anuran distributions to be coupled with variations in environmental energy. For this reason, PETSD was preferred to NPPSD.

Table 3.3: Bivariate correlation matrix for candidate environmental variables. Those variables that were highly correlated (|r| > 0.7), as suggested by Fortin & Dale (2005) and removed from further analysis, are indicated with an asterisk*.

Variables	MAP	MAT	TRange	TSEAS*	PSEAS*	AET*	PET*	NPP	NDV*I	TOPO*	MapSD	MATSD	TRangeSD*	TSEASSD	PSEASSD	AETSD	PETSD	NPPSD*	NDVISD	TOPOSD
MAP	**1**	0.187	-0.534	-0.590	-0.474	0.809	-0.532	0.648	0.884	-0.159	0.530	0.230	0.743	-0.188	-0.302	-0.018	-0.161	-0.109	-0.234	-0.010
MAT	0.187	**1**	-0.375	-0.330	0.234	0.124	0.131	-0.216	0.188	-0.752	0.138	-0.383	0.134	0.081	0.234	0.021	0.084	0.016	0.117	0.006
TRange	-0.534	-0.375	**1**	0.953	0.359	-0.667	0.491	-0.491	-0.453	0.329	-0.246	-0.312	-0.651	0.236	0.295	0.021	0.163	0.107	0.223	0.041
TSEAS	-0.590	-0.330	0.953	**1**	0.499	-0.733	0.563	-0.559	-0.460	0.348	-0.206	-0.289	-0.646	0.277	0.337	0.029	0.189	0.128	0.258	0.030
PSEAS	-0.474	0.234	0.359	0.499	**1**	-0.558	0.494	-0.738	-0.366	-0.019	-0.129	-0.317	-0.442	0.192	0.367	0.022	0.174	0.096	0.247	0.015
AET	0.809	0.124	-0.667	-0.733	-0.558	**1**	-0.700	0.812	0.772	-0.176	0.538	0.351	0.900	-0.233	-0.437	-0.019	-0.215	-0.135	-0.318	-0.029
PET	-0.532	0.131	0.491	0.563	0.494	-0.700	**1**	-0.520	-0.413	0.162	-0.130	-0.383	-0.548	0.655	0.587	0.117	0.356	0.241	0.474	0.026
NPP	0.648	-0.216	-0.491	-0.559	-0.738	0.812	-0.520	**1**	0.662	0.147	0.493	0.383	0.755	-0.153	-0.397	-0.020	-0.181	-0.100	-0.265	-0.028
NDVI	0.884	0.188	-0.453	-0.460	-0.366	0.772	-0.413	0.662	**1**	-0.045	0.781	0.181	0.818	-0.060	-0.246	0.002	-0.121	-0.079	-0.183	-0.016
TOPO	-0.159	-0.752	0.329	0.348	-0.019	-0.176	0.162	0.147	-0.045	**1**	0.032	0.330	-0.105	0.121	-0.044	0.016	0.025	0.056	0.028	-0.019
MapSD	0.530	0.138	-0.246	-0.206	-0.129	0.538	-0.130	0.493	0.781	0.032	**1**	0.114	0.723	0.248	-0.100	0.094	0.001	0.015	-0.037	-0.014
MATSD	0.230	-0.383	-0.312	-0.289	-0.317	0.351	-0.383	0.383	0.181	0.330	0.114	**1**	0.345	-0.090	-0.328	0.027	-0.121	-0.041	-0.191	-0.049
TRangeSD	0.743	0.134	-0.651	-0.646	-0.442	0.900	-0.548	0.755	0.818	-0.105	0.723	0.345	**1**	0.014	-0.423	0.033	-0.166	-0.080	-0.265	-0.038
TSEASSD	-0.188	0.081	0.236	0.277	0.192	-0.233	0.655	-0.153	-0.060	0.121	0.248	-0.090	0.014	**1**	0.318	0.232	0.321	0.268	0.337	0.014
PSEASSD	-0.302	0.234	0.295	0.337	0.367	-0.437	0.587	-0.397	-0.246	-0.044	-0.100	-0.328	-0.423	0.318	**1**	0.152	0.458	0.282	0.513	-0.029
AETSD	-0.018	0.021	0.021	0.029	0.022	-0.019	0.117	-0.020	0.002	0.016	0.094	0.027	0.033	0.232	0.152	**1**	0.463	0.486	0.421	-0.034
PETSD	-0.161	0.084	0.163	0.189	0.174	-0.215	0.356	-0.181	-0.121	0.025	0.001	-0.121	-0.166	0.321	0.458	0.463	**1**	0.888	0.643	-0.018
NPPSD	-0.109	0.016	0.107	0.128	0.096	-0.135	0.241	-0.100	-0.079	0.056	0.015	-0.041	-0.080	0.268	0.282	0.486	0.888	**1**	0.550	-0.015
NDVISD	-0.234	0.117	0.223	0.258	0.247	-0.318	0.474	-0.265	-0.183	0.028	-0.037	-0.191	-0.265	0.337	0.513	0.421	0.643	0.550	**1**	0.047
TOPOSD	-0.010	0.006	0.041	0.030	0.015	-0.029	0.026	-0.028	-0.016	-0.019	-0.014	-0.049	-0.038	0.014	-0.029	-0.034	-0.018	-0.015	0.047	**1**

Table 3.4: Bivariate correlation matrix after removal of transgressor variables.

Variables	MAP	MAT	TRange	PET	NPP	MAPSD	MATSD	TSEASSD	PSEASSD	AETSD	PETSD	NDVISD	TOPOSD
MAP	**1**	0.187	-0.534	-0.532	0.648	0.530	0.230	-0.188	-0.302	-0.018	-0.161	-0.234	-0.010
MAT	0.187	**1**	-0.375	0.131	-0.216	0.138	-0.383	0.081	0.234	0.021	0.084	0.117	0.006
TRange	-0.534	-0.375	**1**	0.491	-0.491	-0.246	-0.312	0.236	0.295	0.021	0.163	0.223	0.041
PET	-0.532	0.131	0.491	**1**	-0.520	-0.130	-0.383	0.655	0.587	0.117	0.356	0.474	0.026
NPP	0.648	-0.216	-0.491	-0.520	**1**	0.493	0.383	-0.153	-0.397	-0.020	-0.181	-0.265	-0.028
MAPSD	0.530	0.138	-0.246	-0.130	0.493	**1**	0.114	0.248	-0.100	0.094	0.001	-0.037	-0.014
MATSD	0.230	-0.383	-0.312	-0.383	0.383	0.114	**1**	-0.090	-0.328	0.027	-0.121	-0.191	-0.049
TSEASSD	-0.188	0.081	0.236	0.655	-0.153	0.248	-0.090	**1**	0.318	0.232	0.321	0.337	0.014
PSEASSD	-0.302	0.234	0.295	0.587	-0.397	-0.100	-0.328	0.318	**1**	0.152	0.458	0.513	-0.029
AETSD	-0.018	0.021	0.021	0.117	-0.020	0.094	0.027	0.232	0.152	**1**	0.463	0.421	-0.034
PETSD	-0.161	0.084	0.163	0.356	-0.181	0.001	-0.121	0.321	0.458	0.463	**1**	0.643	-0.018
NDVISD	-0.234	0.117	0.223	0.474	-0.265	-0.037	-0.191	0.337	0.513	0.421	0.643	**1**	0.047
TOPOSD	-0.010	0.006	0.041	0.026	-0.028	-0.014	-0.049	0.014	-0.029	-0.034	-0.018	0.047	**1**

3.3.1.1 Multicollinearity

The multicollinearity structure of the remaining variables were then analysed using a variance inflation factor (VIF). Multicollinearity occurs when two predictor variables are highly correlated creating a situation in which their effects are difficult to separate (Sykes 2007). In regression analysis, multicollinearity is undesirable because it can increase the standard error scores and the variance of coefficient estimates (Graham 2003).

The VIF score tests for collinear relationships between variables by estimating how much the variance of the estimated regression coefficients are inflated compared to when the predictor variables are not linearly related (Gentleman, Hornik & Parmigiani 2008). In the present study, variables with VIF scores > 10 were considered to be highly collinear (Brien 2007). However, no variables violated this criterion (see Table 3.5) and so the set of variables identified after removing highly correlated variables presented in Table 3.4 were subsequently retained for further analysis.

Table 3.5: Variance Inflation Factor scores of variables retained for modelling.

Variable	VIF
MAP	2.784
MAT	2.806
TRange	2.834
PET	2.049
NPP	4.067
MAPSD	2.096
MATSD	1.737
TSEASSD	2.480
PSEASSD	1.925
AETSD	1.399
PETSD	1.970
NDVISD	2.129
TOPOSD	1.016

3.3.2 **Model selection**

The final set of predictor variables for modelling were derived using a stepwise modelling procedure, specifically forward selection. Stepwise modelling is a semi-automated procedure which builds a model by successively adding variables into the model based on a single or set of predefined criteria (Richerson & Lum 1980). In forward selection, variables are added one at a time into the model. With each addition, the predefined criteria are checked. If the criteria are violated, the variable is omitted from the model (Bini et al. 2009).

For the present study, two selection criteria were set for the forward selection: significance and model fit. Specifically, the included variable had to present a significant contribution towards the model ($p < 0.05$) and increase model fit by at least 1% as determined by the coefficient of determination (R^2). For alpha diversity, three variables satisfied the set criteria: MAP, TRange and NPP. For beta diversity, the final set of variables was the same as of alpha diversity but included MAT and TOPOSD. These variables comprised the final set of predictors used for regression modelling.

3.4 MODELLING

Establishing the statistical relationship between diversity and environmental processes is not straightforward mainly because specifying the correct null model is challenging (Koleff, Lennon & Gaston 2003). The challenge arises from the nature of parametric statistics which assume a lack of dependence between samples or considered data when inferences are made (Legendre 2008). In reality, however, this is not true or likely to be untrue because ecological and environmental processes are spatially structured and hence exhibit varying degrees of spatial dependence and spatial autocorrelation. Spatial autocorrelation is usually governed by the first law of geography: "Everything is related, but near things are more related than distant things" (Charlton & Fotheringham 2009) and refers to the observed pattern in which spatially near locations are more likely to have similar magnitudes for a given phenomenon. For parametric statistics, spatial autocorrelation increases the chance of type 1 error: rejection of the null hypothesis when it is true.

Ultimately, Multiple Linear Regression (MLR) analysis was used to examine the relationship between diversity and the environment. MLR is a statistical procedure which describes the relationship between a single response (y) and multiple predictor (x_0 - x_n) variables using a line of best fit. Regression analysis was selected because it is an established statistical technique and has been widely used in spatial analysis (Cade & Noon 2003; Diniz-filho et al.

2007; Qian et al. 2007; Wang 2013; Murphy et al. 2015). Regression analysis constitutes the raw data approach to modelling diversity as envisaged by Legendre et al. (2005).

Two regression models were used to examine the relationship between diversity and the environment, one global model, the Ordinary Least Squares (OLS) model and one local, the Geographically Weighted Regression (GWR). Global regression methods (OLS) assume that the relationship under study is constant (i.e. stationarity), and therefore the estimated parameters remain constant in space, which is not always true. In contrast, the GWR allows for local influences, leaving the parameters free to vary spatially and fit more appropriately. Although the technique does not allow extrapolation beyond the study region, by allowing the parameters to vary locally within the study area a more accurate and appropriate basis for descriptive and predictive purposes is achieved (Foody 2004).

The high number of species with small geographic ranges and the geographic complexity of study area could produce variations in the relationship among diversity and climate in different regions the study extent. A global model with a single best predictor will be the best choice for extrapolation or to know the general relation between variables under study. However, to determinate how relationships between predictor and response variables varies in the space, a locally weighted model can help to reveal spatial variation in the empirical relationships between variables that otherwise might be ignored in the analysis. For this reason, both a global a local method was implemented.

3.4.1 **Global method: Ordinary Least Squares regression (OLS)**

The first regression model employed to relate alpha diversity to environmental processes was an Ordinary Least Squares (OLS) regression. OLS is the standard linear regression model and can been viewed as (Equation 3.3):

$$Y = x\beta + \varepsilon$$

Equation 3.3: OLS model components.

Where Y is the response variable, x is the predictive variable, β is the estimated parameters or constants, and ε is the vector containing the error terms (Ortiz-Yusty, Páez & Zapata 2013).

For an OLS regression with multiple independent variables, the regression model is presented as (Equation 3.4):

$$Y = \beta_0 + \beta_1 X_1 \ldots \beta_n X_n + \varepsilon$$

Equation 3.4: Standard form of multiple predictor OLS.

where *Y*, is the dependent variable, x_1 to x_n are the explanatory variables, β_0 is the regression intercept, β_1 to β_n are estimated coefficients for the *n* number of independent variables, and ε is the error term of the model (Winship 1984; Di Virgilio et al. 2014).

To make valid statistical inferences, the OLS model follows a discrete set of assumptions, which include:

1. Residuals to be independent or uncorrelated; and
2. Response and predictors variables must exhibit stationary relationship across data space

Spatial and ecological data often violate assumptions one and two because spatially structured variables are often auto-correlated by nature. Hence, the OLS model is not recommended in studies that include spatially distributed variables (Dormann et al. 2007).

Accordingly, the spatial autocorrelation structure of the OLS model residuals was tested using a global Moran's I function, which ranges from –1 for complete negative spatial autocorrelation to +1 for complete spatial autocorrelation with 0 indicating the absence of spatial autocorrelation. For both alpha and beta diversity, OLS residuals were found to be significantly ($p < 0.05$) auto-correlated (Moran's $I > 0.85$). For this reason, the OLS model was used for comparative purposes only.

3.4.2 **Local method: Geographically Weighted Regression (GWR)**

Global regression models, such as OLS, assume that the relationships being studied are the same everywhere within the study area (Czarnota, Wheeler & Gennings 2015). However, species-environment relationships are seldom homogenous, assuming spatially variable relations across geographic extents (Foody 2004). This phenomenon is known as spatial nonstationary and comprises the second order effect in geographical analysis.

Empirically, spatial nonstationarity is a form of model misspecification resulting from intrinsic local differences or variations in pattern process relationships across the study extent. In the presence of nonstationary relationships, results from global models may limited in their descriptive and predictive power (Sheehan, Strager & Welsh 2013).

To assess whether non-stationary relationships were present across the study extent, the residual structure of OLS models were inspected using the Koenker Bruesch-Pagan (BP) statistic. The Koenker BP statistic tests for spatial nonstationarity by determining whether the predictor variables in the model have a constant relationship with the response variable in

geographic and data space. If the Koenker score is significant, the modelled relations are spatially variable. For both alpha and beta diversity, Koenker scores for OLS residuals were found to be significant ($p < 0.05$) indicating a heteroskedastic residual structure and the presence of spatial nonstationarity.

For this reason, a local regression model, the Geographically Weighted Regression (GWR), was used to account for the inconsistent spatial relationships. GWR is refinement of the standard linear model which computes a regression equation for each feature in the study area thereby allowing parameters to vary continuously in geographic space leaving them free to fit more appropriately (Dormann et al. 2007; Nthiwa 2011). It differs from global models in not looking for local variation in 'data' space, but by moving a weighted window over the data across the entire study area, estimating one set of coefficient values at every chosen 'fit' point (Dormann et al. 2007; Tsai 2011). The estimation of the GWR is given by (Equation 3.5):

$$Y = \boldsymbol{\beta_0}(\boldsymbol{u},\boldsymbol{v}) + \boldsymbol{\beta_1}(\boldsymbol{u},\boldsymbol{v})\boldsymbol{x_1} + \ldots + \boldsymbol{\beta_u}(\boldsymbol{u},\boldsymbol{v})\boldsymbol{x_n} + \boldsymbol{\varepsilon}$$

Equation 3.5: Standard form of the GWR.

Where β_0 is the regression intercept, $\beta_1 - \beta_u$ are the coefficients intercepts of the independent variables, $x_1 - x_n$ are the independent variables and (u, v) are the spatial coordinates of each location in the study extent and ε is the error term. For the GWR, nearer observations have a greater weight than distant ones in calibrating the local set of regression parameters (Czarnota, Wheeler & Gennings 2015). The relationship about each point *i* is measured using a weighted least squares approach and is given by the expression (Equation 3.6):

$$\widehat{\beta}_i = (X^T W(u_i, v_i) X)^{-1} X^T W(u_i, v_i) y$$

Equation 3.6: Weighted least squares estimation.

Where X is the matrix of independent variables, y is the dependent variable, $W(u, v)$ is the weighting matrix of each observation whose diagonal represents the spatial weighting of observation about point i.

In the present study, GWR models were structured in the following manner. First, a spatial weighting kernel was defined, specifically a fixed spatial kernel. The spatial kernel determines how spatial neighbours are weighted. Fotheringham, Brundson & Charlton (2002) notes that fixed kernels are appropriate if the sampling units or points under study are equal in size or constant in space such as the grid system used in the present study. Next, a neighbourhood search radius (bandwidth) was chosen for the GWR. The model bandwidth determines how many neighbours are weighted in deriving parameters for a local focal

location. Smaller bandwidths are preferred to larger ones because as the size of the bandwidth is increased the parameters of the GWR strive towards those of global OLS model. Initially 30 neighbours were selected, which is the default in ArcGIS, and were then sequentially increased by 10 neighbours until the model was stable and errors in the output attribute table were removed. In the end, 50 neighbours were used as the model bandwidth which corresponds to a radius of ±407 km.

In the GWR, both untransformed and $\log_{10}$ transformed versions of the predictor variables were used, but the untransformed data consistently performed better than the log data, therefore it was decided to only present the untransformed data for the GWR.

3.4.3 **Model performance**

Selection procedures which allow for the identification of the best model or sets of best models to describe spatial patterns under study are lacking in most ecological and geographical studies where data are spatially autocorrelated (Dole et al. 2012). Kissling & Carl (2008) introduced a conceptual framework for evaluating and identifying best performing regression models in the presence of spatial autocorrelation. The framework is based on three idealised model parameters which the model must satisfy: **(1)** To maximise model fit so as to make more accurate predictions, **(2)** efficiently reduce the spatial autocorrelation in the residuals to avoid type I error and **(3)** present the best model parsimony, that is, to account best for model complexity with respect to the number of variables and variance explained.

To identify the best performing model(s) the present study adopted the above-mentioned approach when evaluating regression models. The parameters used to evaluate each criterion were:

1. **The adjusted coefficient of determination for model fit (Adj. R^2).**

   Where the adjusted R^2 is a measure of model performance on a 0-1 scale and indicates how much variation in the response variable is explained by the predictor variables. The higher the adj. R^2 value of the model, the better the fit presented by the model (Kissling & Carl 2008).

2. **The Moran's I value for spatial autocorrelation.**

Moran's I values closer to 0 will indicate a sufficient reduction in spatial autocorrelation and a better performing model with respect to spatial dependence (Fortin et al. 2012).

3. **The Akaike Information Criterion for model performance (AIC).**

   Based on information theory, the AIC is a measure of model quality with a lower AIC score indicating a better performing model (Kissling & Carl 2008; Fischer & Getis 2010). The AIC can be used to compare the performance of models as long as the response variable between those models remains consistent.

The best performing model was the one which could discriminate the best amongst the above mentioned evaluation criteria. Maps of the GWR models standardised residuals, coefficient slopes, local r^2 and conditions numbers (CN) were produced to evaluate the performance of the GWR and to observe spatially variable relationships across the study extent. Standardised or residuals are residuals divided by an estimate of its standard deviation. Standardised residuals (stundentized) help in accounting for the different variances of the residuals and are useful for the detection of outliers (i.e. to detect where the model is misspecified). The local r^2 indicates the model fit in specific locations of the study area while conditional numbers (CN) indicate local collinear relationships. For comparative purposes, maps of the standardised residuals for the OLS models were also created.

3.5 RESULTS

3.5.1 Patterns of alpha diversity

Figure 3.8 shows the spatial distribution of alpha diversity across extent. Shades of red indicate high alpha diversity and shades of green indicate areas of low alpha diversity. The hue of each colour expresses the intensity of diversity (i.e. light green > dark green). Mean alpha diversity (γ-diversity) was 26 species (±15 SD species).

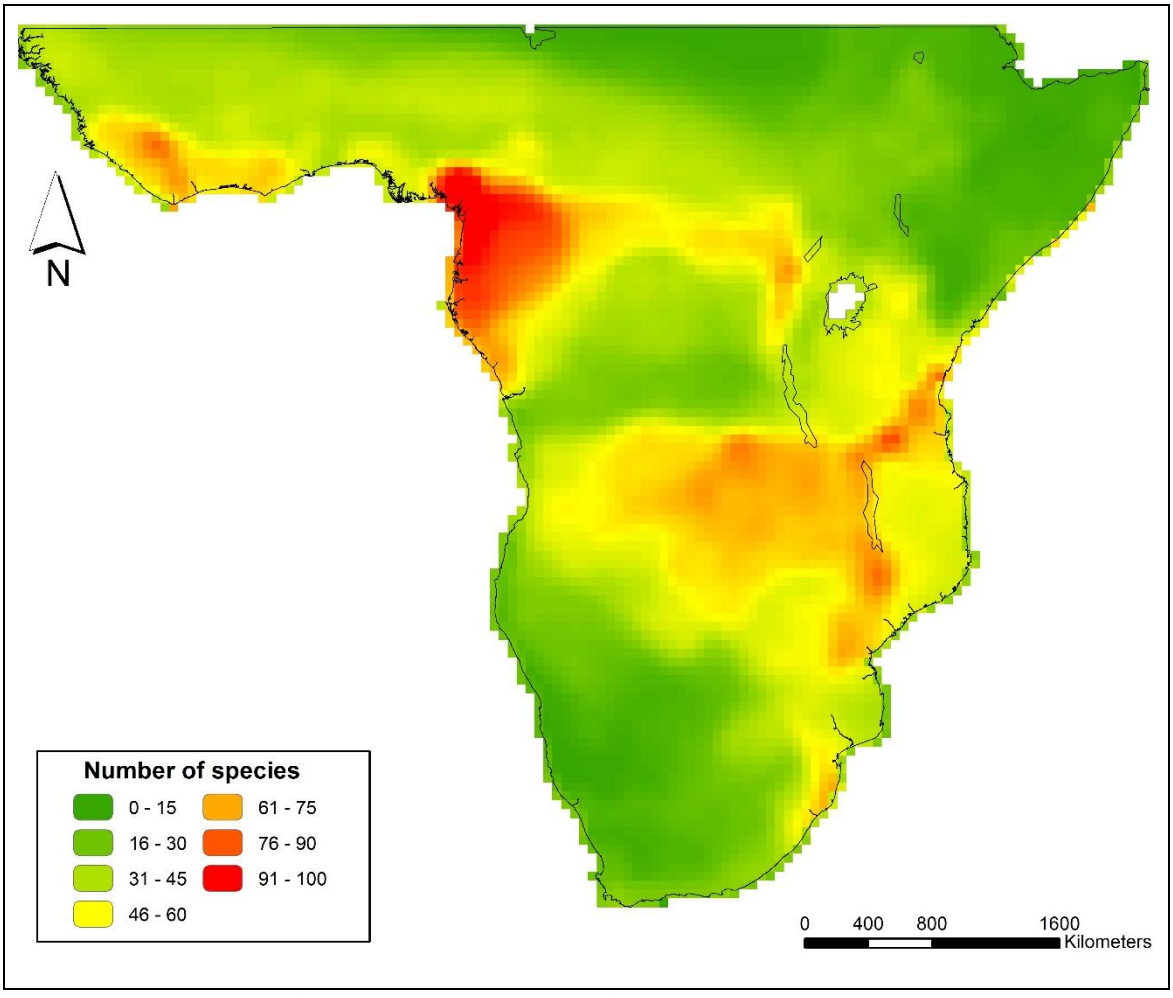


Figure 3.8: Spatial distribution of anuran alpha diversity across the Afrotropics at 0.5º × 0.5 º resolutions as measured per zeta diversity partitioning using the first order component (ζ_1).

The spatial distribution of alpha diversity was highly variable across extent. Tropical Africa exhibited the highest levels of alpha diversity (> 70 species per quadrat). The northern Congo basin, the coast and adjacent interior east of the Gulf of Guinea from southern Gabon through to the Niger Delta were particularly species rich. The Cameroonian Highland Forest complex north-east of Mount Cameroon, north of Mount Kupe towards Mount Manengouba (Figure 3.9) were the most species rich area across the Afrotropics with a mean alpha diversity four times higher than the regional mean (± 100 species). High levels of alpha diversity were also

observed in West Africa near the Guinean forest-savanna complex from northern Senegal south through eastern Liberia. In East Africa, the Eastern Arc Mountains of Tanzania extending from the Pare Mts. in the north, south towards the Udzungwa Mts near the Malawi diversity areas of high local diversity.

Intermediate alpha diversity*[3] (31 - 61 species) was observed across much of South-eastern Africa. In general, alpha diversity was higher in moist regions and attenuated towards latitudinal extremes with the arid Sahel and Namib – Kalahari Desert regions recording the lowest number of species (< 15).

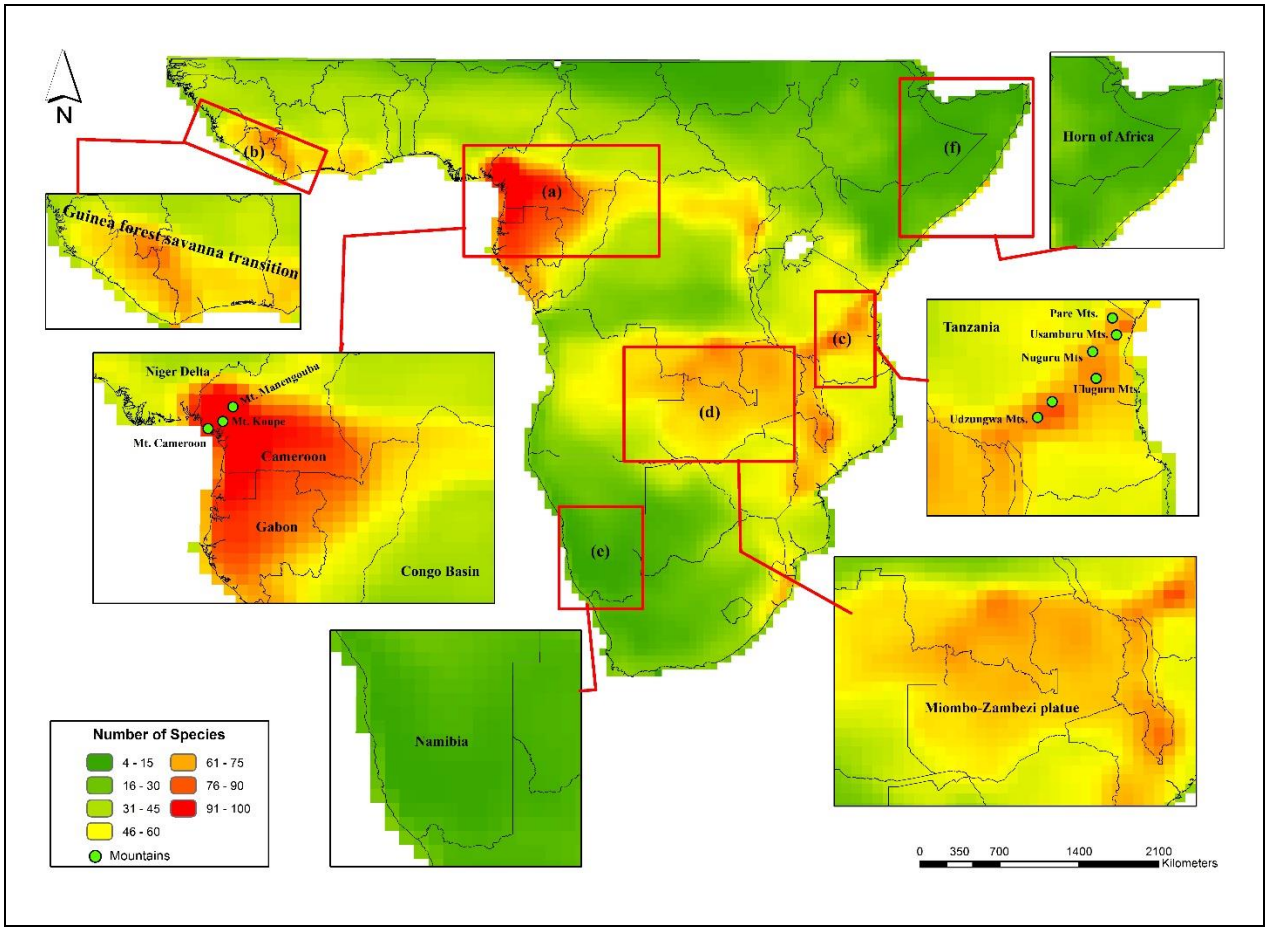


Figure 3.9: Areas of high (a-c), intermediate (d) and low (e-f) alpha diversity. Tropical Africa (a) through the Gulf of Guinea, Cameroon and the northern Congo basin area had the highest alpha diversity. West African along the Guinea savannah-forest transition (b) and east Africa along the Eastern Arc Mountains (c) were areas also areas of high diversity. The Miombo woodlands region (d) presented intermediate diversity. Arid regions such as the Namib Desert (e) and Horn of Africa (f) were areas of low richness.

The geographic variation in anuran alpha diversity presented here are concordant with established literature (Buckley & Jetz 2007; Qian 2009, 2010). The consensus is that: anuran richness across the Afrotropics increases towards the equator, attaining concentrations highest in the moisture saturated areas of the tropics and subtropics, and then attenuates towards arid regions of the subcontinent. This observation has been documented to persist for

[3] Based on dividing the species richness distribution histogram into thirds.

all orders of amphibians, across a range of spatial scales and across different geographical realms (Seymour et al. 2001; Burgess et al. 2002; Buckley & Jetz 2007; Pineda & Lobo 2009; Qian 2010). The high anuran diversity observed across much of Tropical Africa and moisture-saturated areas may be owed to the high availability of resources here which are needed for survival and reproduction. Findings by Hawkins et al. (2003) seem to supports this notion, linking global vertebrate diversity to areas of high energy (water and ambient energy) and productivity. Furthermore, the exceptional high diversity documented in Cameroon into Gabon through the northern Congo Basin may reflect the historical influence of climate on diversity in this region. For example, the Cameroonian Highlands comprise alternating slopes of extinct volcanic cones which have resulted in the isolation and speciation of some species through dispersal based processes (i.e. allopatric speciation). In addition, this region has been climatically more stable than the southern parts of the Congo Basin which have been prone to frequent climatic shifts resulting in ‘savannafication’ (Oslisly et al. 2013). The older assemblage structure would exert support to the climate stability hypothesis which posits that older assemblages have more species because speciation events accumulate over time. This pattern is consistent to that of Baselga, Gomez-Rodriguez & Lobo (2012) who documented that tropical areas harboured more amphibian species owing to more stable climatic conditions typically associated with low latitudes.

3.5.2 **Patterns of beta diversity**

Figure 3.10 shows the spatial distribution of beta diversity as turnover across the study extent. Shades of red indicate high biotic dissimilarity and shades of green indicate high biotic similarity. The hue of each shade indicates the relative intensity of the dissimilarity.

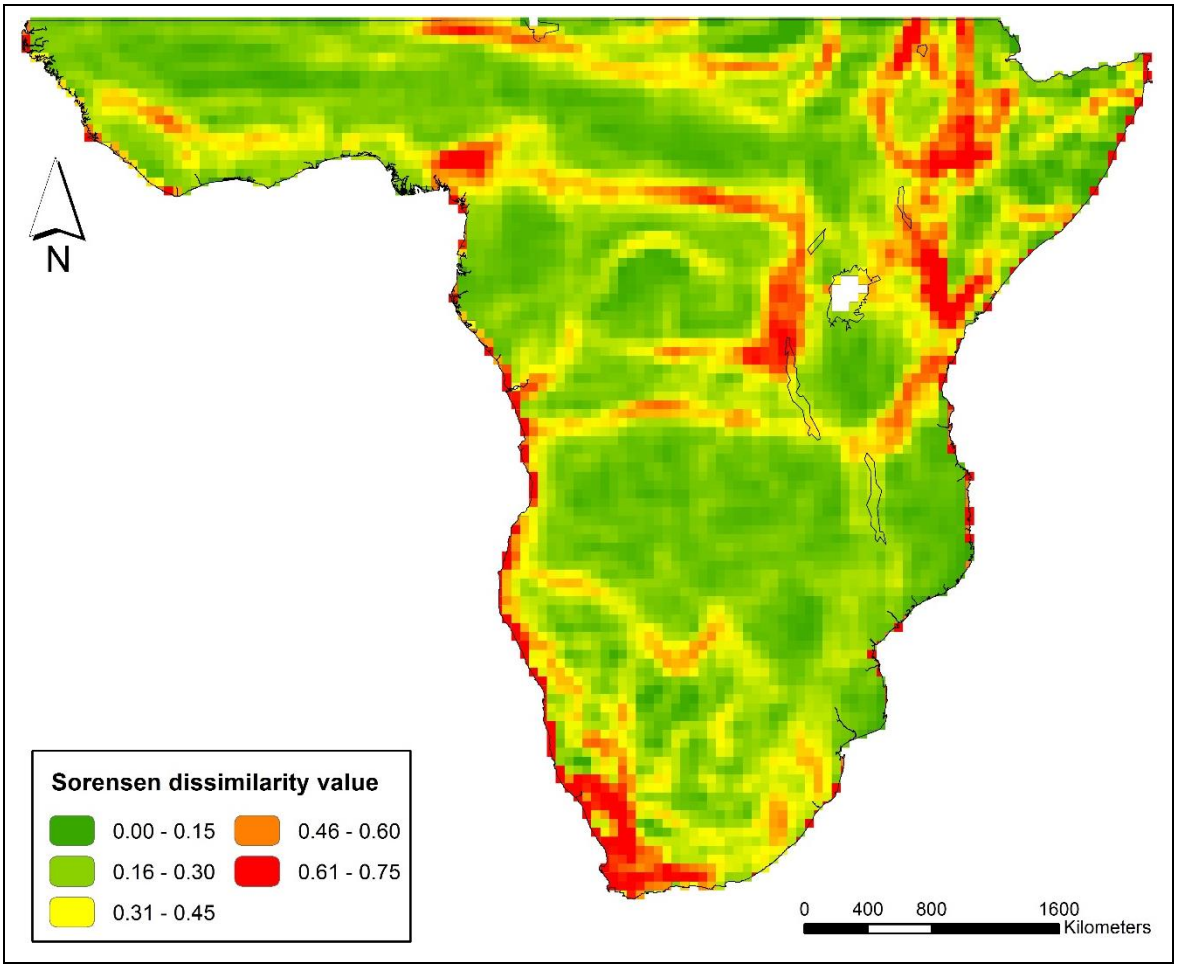


Figure 3.10: Spatial distribution of beta diversity across the Afrotropics at 0.5º × 0.5º resolution as using the Sorensen index.

In contrast to patterns of species richness, where diversity is concentrated in tropics, high levels of beta diversity occurred across a range of latitudes. East Africa, from the Ethiopian Highlands through the Congolian – Albertine Montane Forest Complex, were areas of high beta diversity (Figure 3.11a,b). High beta diversity was also observed in north-central Africa distributed in a front like manner along the Northern Congolian Forest-Savanna Mosaic border west through the Cameroonian highlands to the Guinean Montane Forest complex (Figure 3.11c,f). In central Africa, specifically the Congo Basin, regions flanking the Congo river also presented high beta diversity (β_{sor} = ±0.4). High beta diversity observed along the coastline of study extent was assumed to be due to sampling bias.

The Miombo Woodlands/Congo forest transition presented intermediate beta diversity. In contrast, beta diversity was generally low across much of Southern Africa, with the exception

of the Cape Floristic region north through the Karoo region, in South Africa, where high beta diversity, equivalent to the maximum, was recorded (β_{sor} = 0.7) (Figure 3.11e). Low beta diversity (β_{sor} < 0.1) was also observed across much of the northern Afrotropics along the Sahel region and South-eastern Africa across the Zambezi plateau. Broadly, many areas of high beta diversity coincided with known biogeographical transitions, while low beta diversity persisted within biogeographical provinces.

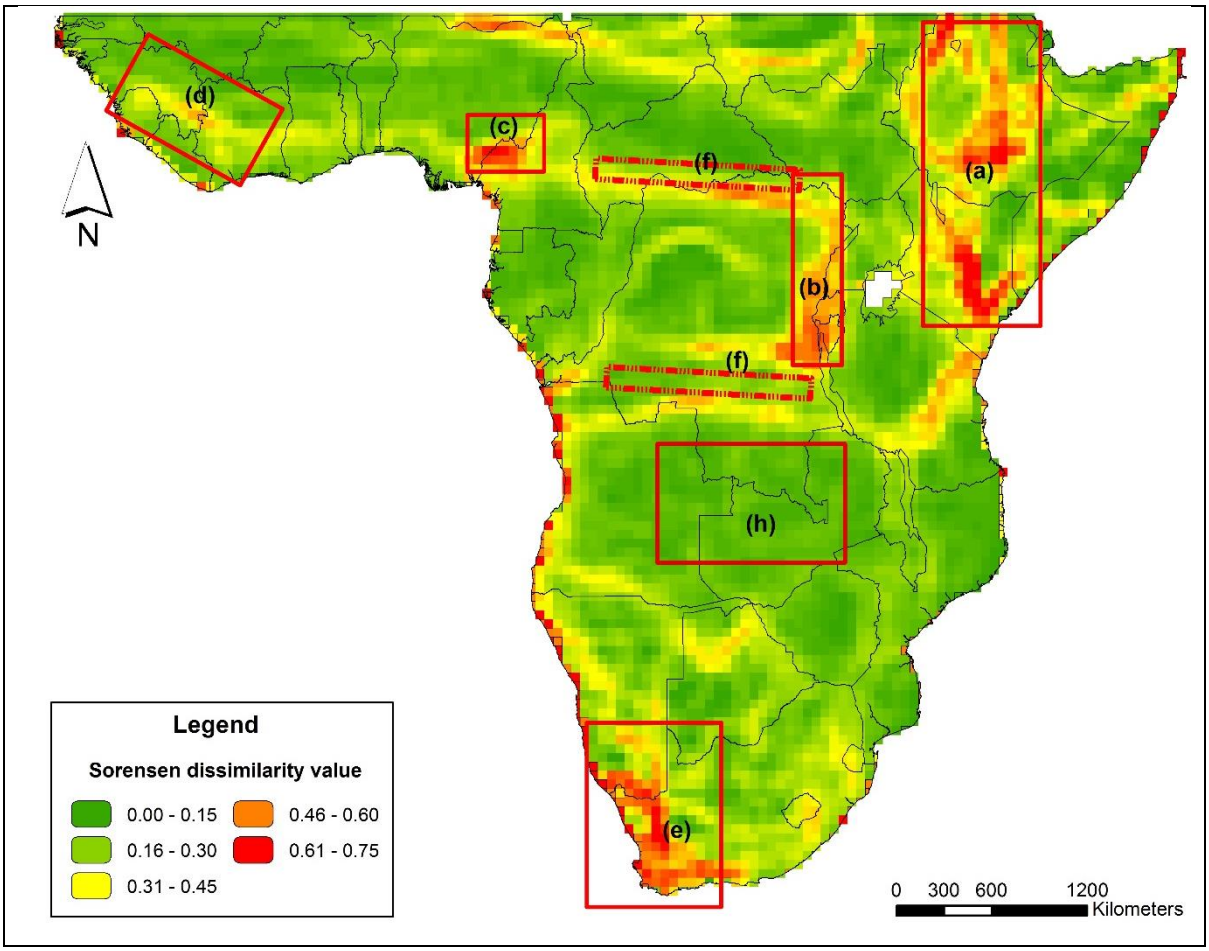


Figure 3.11: Regions of high (a-g) and low (h) beta diversity across the Afrotropics.

The patterns of beta diversity presented here deviate from those of Baselga, Gomez-Rodriguez & Lobo (2012). Discrepancies in the spatial distribution of beta diversity between the present study and Baselga are surprising since both studies incorporate the Sorensen index to document beta diversity. Discordant results may be owed in part to the resolution of grid cells at which patterns were documented (0.5º vs 1º), although Melo, Rangel & Diniz-Filho (2009) found beta diversity patterns to remain consistent under altered grain size, and the systematic quantification of beta diversity between studies. Specifically, in the present study, a neighbourhood analysis was used to derive spatial distributions of beta diversity per zeta diversity partitioning whereas Baselga, Gomez-Rodriguez & Lobo (2012) used an intra-

cell approach to quantify biotic heterogeneity using a modified version of the Sorensen index. The findings presented here, however, do corroborate those patterns documented by Linder et al. (2012). The concordant spatial patterns reflect concordant spatial analysis: both studies employed neighbourhood analysis to generate patterns of biotic heterogeneity though the indices used to quantify these were different (β_{sim} vs β_{sor}). Despite this, the emergent patterns were similar. Both studies found the eastern and northern Congo basin, central Cameroon through West Africa and the Cape Floristic region as areas of high beta diversity. High beta diversity across these regions may in part be owed to the large number of narrow ranged and endemic species encountered here (Seymour et al. 2001). Narrow range species express lower environmental tolerances and thus disperse less readily than wide ranged ones. Martin & Ferrer (2015) recently linked the low species richness and high spatial turnover of amphibians associated with Mediterranean biomes, similar to those of the Cape region, to range size and evolutionary history. In particular, they note that Mediterranean regions are areas of high intra-annual climatic variability which have resulted in greater degrees of specialisation amongst taxa inhabiting these regions.

Furthermore, the high beta diversity in East Africa may be owed to the complex geologic conditions present (uplift, crustal tear) here which have sculpted highly heterogeneous environments resulting in high diversification through discrete evolutionary responses. The tendency of high beta diversity to coincide with biogeographic transition zones or fringes may too reflect discrete responses of species to complex climatic and topographical patterns present here which result in dynamic biogeographical signals. These findings support those of Van Rensburg et al. (2004) who correlated the turnover of South African birds with biome transitions. Specifically, that biogeographic transitions (Araújo 1996; Williams, De Klerk & Crowe 1999; Aronson et al. 2002; Di Virgilio, Laffan & Ebach 2012) are areas of high environmental heterogeneity characterised by complex transitions of different flora from different biogeographic regions and thus furnish more niches and hence have greater degrees of species filtering (Araújo 1996; Williams, De Klerk & Crowe 1999; Aronson et al. 2002; Di Virgilio, Laffacn & Ebach 2012). Furthermore, both studies found low biotic heterogeneity remains confined to the interior of biogeographical provinces (i.e. Congo basin, Miombo plateau) thereby asserting support to the climatic stability hypothesis, which postulates that regional species pools are at equilibrium where climatic conditions are more stable (Fjeldsa, Lambin & Ivlertens 1999; Diniz-filho, Diniz-filho, Barberi & Lima-riberio 2010)

3.5.3 **Determinants of alpha diversity**

3.5.3.1 OLS

With the OLS model, significant relationships ($p < 0.05$) between alpha diversity and environmental correlates were obtained (Table 3.6). However, because of the significant Koenker (BP) scores associated with the OLS model, the robust probability associated with each environmental variable were inspected to determine whether Pearson probabilities were valid. Based on the significant robust probabilities ($p < 0.05$), all variables presented a significant contribution to accounting for variation in amphibian species richness. Coefficient slopes for all predictors presented positive slopes with MAP and NPP presenting the steepest slopes. In terms of model fit, the OLS indicates that environmental variables could explain 62% of the variation in alpha diversity, which is similar to the results obtained by Seymour et al. (2001) for the anurans of Southern Africa.

Table 3.6: OLS diagnostics obtained for alpha diversity.

	Diagnostic			
	Coefficient	**Robust t- value**	**p - value**	**Robust p - value**
Intercept	-1.15	27.59	< 0.05	< 0.05
MAP	0.16	27.73	< 0.05	< 0.05
TRange	0.02	3.87	< 0.05	< 0.05
NPP	0.15	36.04	< 0.05	< 0.05
Model Performance		**z-value**		
Adj. R2	0,62			
AICc	-3241.89			
Morans I	0,92	99.94	< 0.05	
Koenker	180.74		< 0.05	
Jacque Bera	332.881		< 0.05	

Furthermore, results from the explanatory regression (forward) showed that MAP was the single strongest predictor of anuran diversity followed by NPP. MAP could explain 51% of the variation in anuran diversity while NPP, as a secondary determinant, explained 41%, of the variation. TRange was the weakest predictor explaining only 8% of the variation in anuran diversity. The importance of water (MAP) and the interplay between water and temperature to produce productive environments (NPP) in constraining amphibian diversity documented here, correspond to those findings of Buckley & Jetz (2007) for amphibians at a global scale.

Potential bias in the OLS model was then determined by inspecting the Jacque-Bera statistic associated with the model residuals. The Jacque-Bera statistic, which checks model accuracy by testing normality of the residuals, was significant ($p < 0.05$) indicating the presence of model bias. To source the bias, the standardised residual map (Figure 3.12), residual histogram (Figure 3.13a) and scatter plot (Figure 3.13b), depicting observed to predicted values, produced by the model were inspected. Inspecting the standardised residuals maps of the OLS regression showed that the model consistently over predicted in areas of high diversity (shades of red) and under predicted in areas of low diversity (shades of blue).

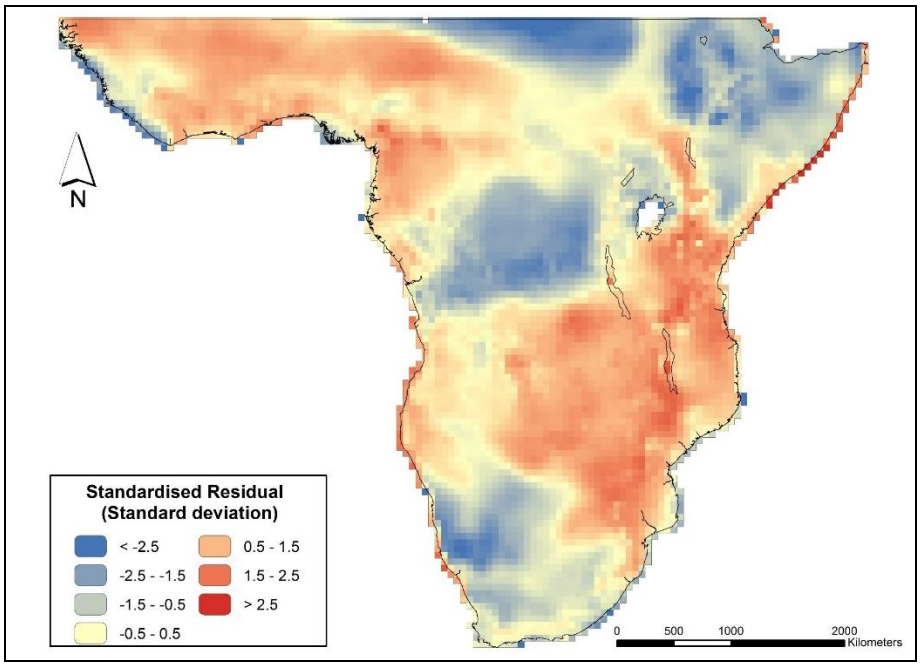


Figure 3.12: Spatial distribution of standardised residuals produced by the OLS model across the study extent. Shades of red indicates where the model overestimates while shades of blue indicates where the model underestimates.

Considering the significant and high Morans I score ($I = 0.92$, $z = 99.94$, $p < 0.05$) associated with the OLS residuals (Figure 3.13a), model bias was most likely due to misspecification, specifically omitted or missing predictor variables.

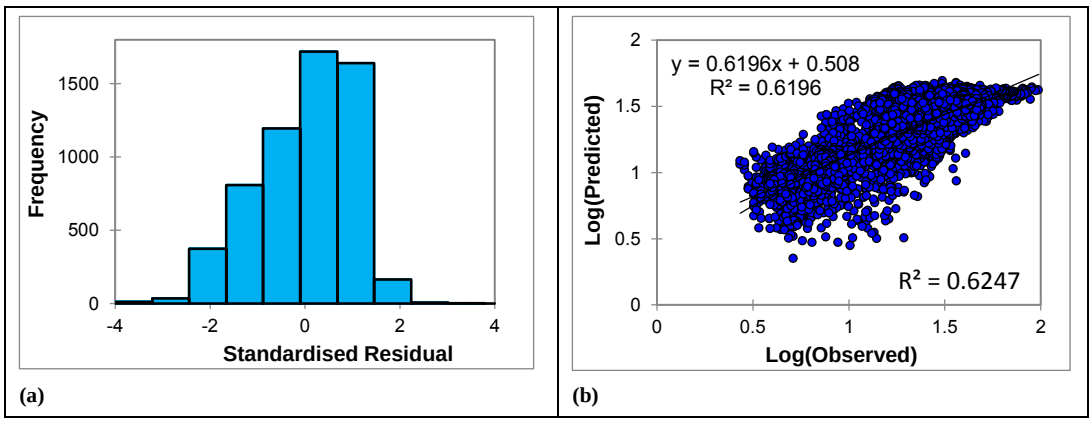


Figure 3.13: (a) Histogram showing the distribution of model residual. (b) Scatter plot of predicted to observed model values with fitted linear and nonlinear trend lines. Model misspecification is indicated by the improved fit of the nonlinear ($R^2 = 0.625$) to the linear line ($R^2 = 0.619$).

However, fitting linear and nonlinear trend lines to the scatter diagram of the predicted to observed values of the OLS model showed that misspecification could also be attributed to nonlinear interactions between the response and predictor variables as suggested by the improved model fit offered by fitting a nonlinear trend line to the data (R^2 = 0.625) (Figure 3.13b). Using a nonlinear or quadratic function to transform the data may thus have improved model performance, though this improvement may have been negligible as suggested by the minimal increase in the R^2 for the nonlinear trend line (R^2 = 0.624).

3.5.3.2 GWR

With the local GWR (Table 3.7), all predictors presented significant contributions towards the model ($p < 0.05$). The median coefficient slopes presented by the predictors in the GWR were similar direction and magnitude to those of the OLS.

Table 3.7: GWR model parameters.

	Diagnostic		
	Coefficient		**p - value**
Intercept	-1.49		< 0.05
MAP	0.16		< 0.05
TRANGE	0.02		< 0.05
NPP	0.15		< 0.05
Model Performance		**z-score**	
Adj. R^2	0.92		
AICc	-10092.12		
Morans I	0.62	68.08	< 0.05

The relationship between model parameters, model coefficients and predictor coefficients were spatially variable across the extent (Figure 3.14).

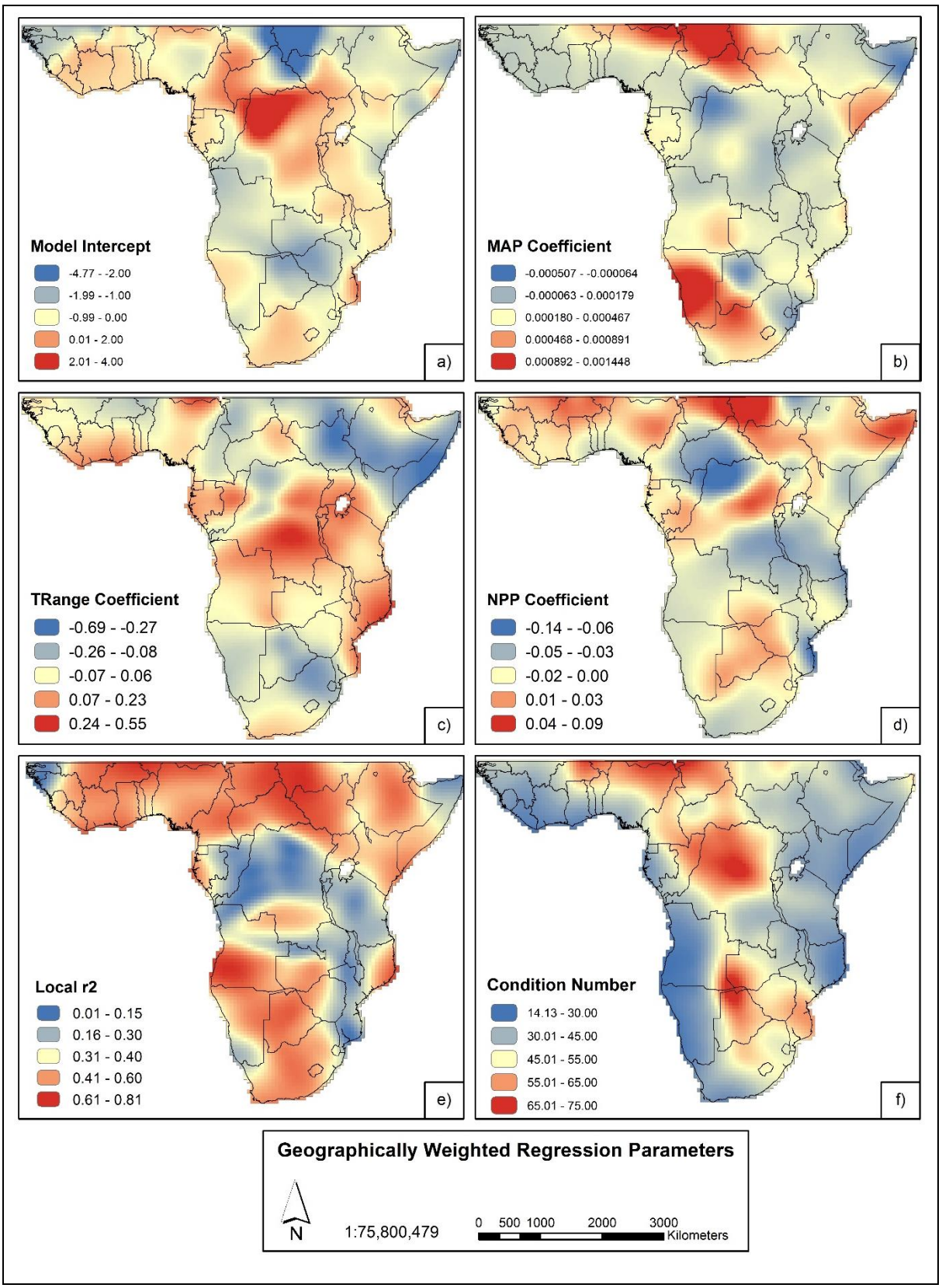


Figure 3.14: Spatial variation in (a) model intercept, (b-d) predictor coefficients, (e) local r^2 and (f) condition numbers across the study extent at 0.5° × 0.5° resolution for alpha diversity derived by the GWR model.

Simple interpretation of the spatial distribution of model parameters were not straight forward, but seemed to indicate the following: MAP coefficients assume positive and steep slopes in arid regions attenuating and becoming negative towards moist regions (Figure 3.14b). TRange coefficients slopes were positive and steep in warm moist areas and more gentle and negative in

cooler, climatically variable areas (Figure 3.14c). For NPP (Figure 3.14d), coefficients seem to be spatially random although steep positive slopes persisted in West and Southern Africa while negative slopes persisted across much of the central and eastern portions of study extent. With the local r^2, which indicates local model fit (Figure 3.14e), the highest fit was obtained across the northern and southern Afrotropics while poor model fit (low r^2) was obtained in the tropical and subtropical regions from the Congo Basin across the Miombo Woodlands in Tanzania south into Mozambique. Low fit in the tropical latitudes may indicate the omission or exclusion of a key explanatory variable in these regions. Such a metric, as suggested by Hawkins et al. (2003), may pertain to either energy availability, which limits vertebrate diversity in tropical regions, or the ecological history of the taxa or region which not explicitly considered in this study. Furthermore, high local condition numbers (CN > 30), notably those across the central study extent, show that local collinearity was still present amongst the data. Consequently, coefficients for these localities may be biased (Figure 3.14e).

Though significant, the lower Morans I ($I = 0.62$, $z = 68.08$, $p < 0.05$) value associated with the GWR showed that the model accounted better for spatial dependence amongst the residuals than the OLS model. However, inspecting the standardised residual map (Figure 3.15) of the GWR showed that like the OLS, though to a lesser degree, the GWR model consistently overestimated in areas of high species richness and underestimated in areas of low species richness. This significant autocorrelation reaffirmed model misspecification by missing predictor variables in these regions.

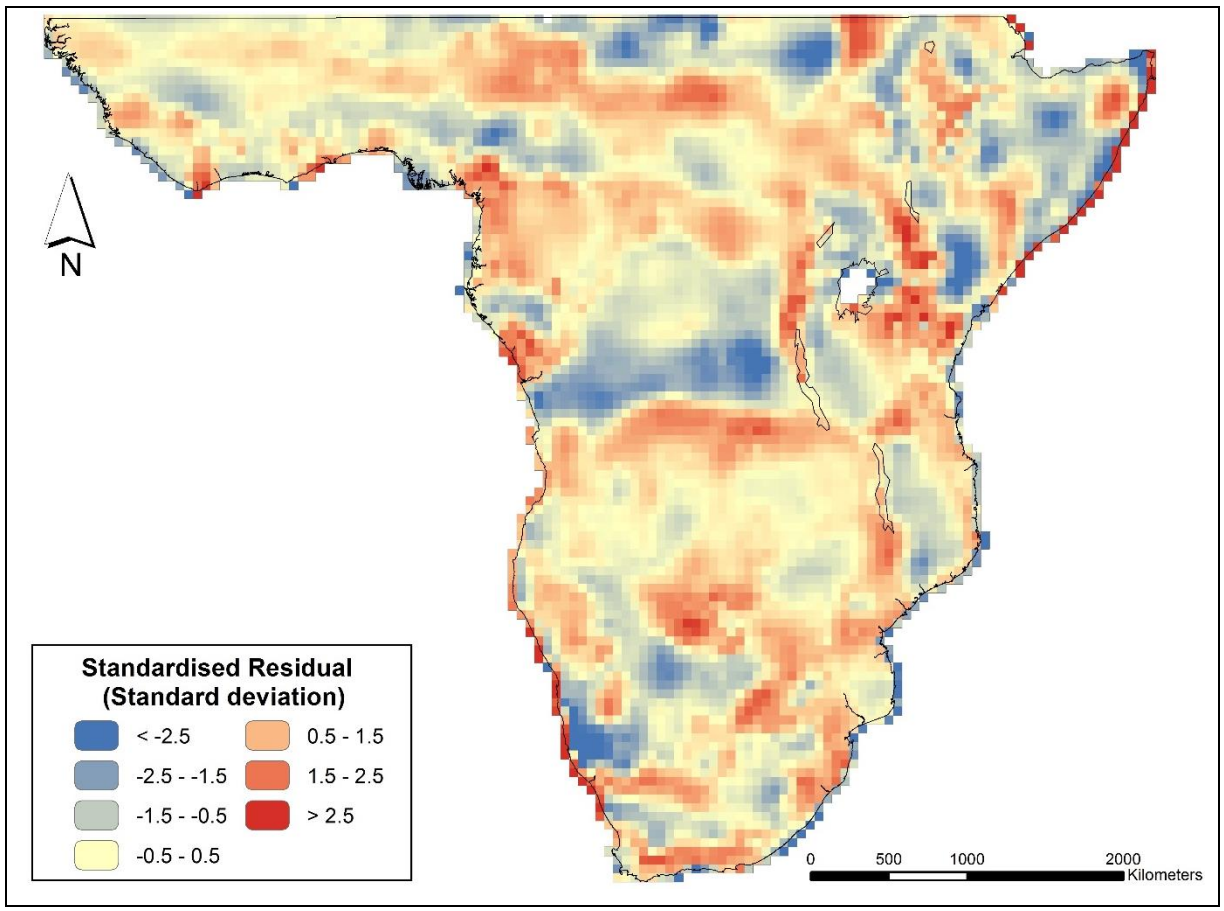


Figure 3.15: Spatial distribution of standard residuals produced by the GWR model for alpha diversity across the study extent. The model overestimates in areas of high alpha diversity (red) and underestimates in areas of low species richness (blue).

Overall, the GWR could explain 92% of the variation in anuran diversity which is a significant improvement over the 62% accounted for by the OLS model (Figure 3.16a). The improved fit offered by the GWR model is demonstrated by the observed to the predicted scatter diagram presented in Figure 3.16a. The frequency histogram (Figure 3.16b) further shows how the GWR better corrected for spatial autocorrelation than the OLS by reducing standard residual scores such that a less skewed distribution was achieved.

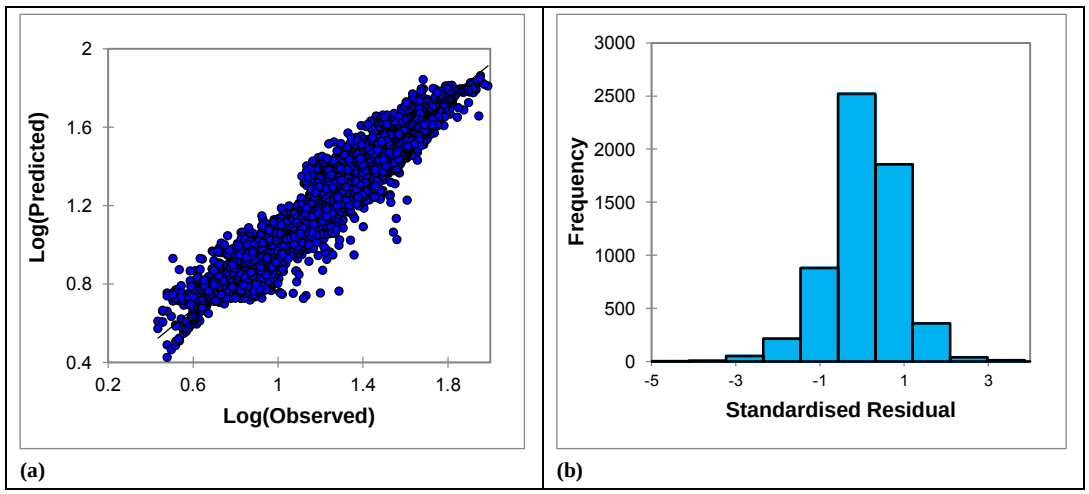


Figure 3.16: (a) Scatter plot for predicted and observed values of GWR models with corresponding (b) histogram of standardised residuals.

3.5.3.3 Model Comparison: Overview

When comparing the models, the GWR model recorded the lowest AIC score (Table 3.8), which indicates that the model had a better balance between the amount of variance explained and the number of variables included.

Table 3.8: Comparison of OLS and GWR performance for alpha diveristy.

Model	Adj. R^2	AIC	Morans I	z- score	p-value
OLS	0.62	-3241.89	0.92	99.94	< 0.05
GWR	0.92	-10092.12	0.62	68.08	< 0.05

The GWR also corrected best for spatial dependence in model residuals as indicated by the significantly lower Moran's I scores (GWR = 0.62, OLS = 0.92). The improved performance of the GWR in accounting for spatial autocorrelation has been attributed to its ability to allow for spatially variable relationships and the consideration of neighbour influence during model calibration (Propastin et al. 2008; Sheehan, Strager & Welsh 2013). Furthermore, the GWR also explained the most variation in amphibian richness; specifically the GWR could explain 92% to the 62% accounted for by the OLS. When considering the model evaluation criteria, it can be deduced that the GWR consistently outperformed the OLS model in all criteria further showing the data were better suited to a local and not a global model. The improved performance of the local GWR to global OLS documented here are concordant with several other published studies (Fischer & Getis 2010; Ortiz-Yusty, Páez & Zapata 2013; Sheehan, Strager & Welsh 2013).

The results thus show that parameters and estimates derived from global models fail to capture and represent the spatially complex relationships between alpha diversity and environmental correlates across the study extent. Local models, such as the GWR, thus comprise ideal explanatory tools to supplement and support parameters and findings derived through global methods.

3.5.4 **Comparison to other studies**

Results from regression analysis permit for some discrimination to be made with respect to the various hypotheses regarding the relationship between alpha diversity and contemporary environmental conditions. The regression models suggest that mean contemporary environmental conditions are strong determinants of anuran alpha diversity across the Afrotropics, accounting for up to 92% of the variation in species richness patterns depending on the regression model employed. These findings coincide with those of parallel studies (Buckley & Jetz 2007; Werner et al. 2007; Pineda & Lobo 2009; Ortiz-Yusty, Páez & Zapata 2013; Qian & Wang 2015). This asserts further support to the species-environment hypothesis and reaffirms

the importance of contemporary environmental conditions in establishing present species distributions of alpha diversity.

The substantive predictive power of mean annual precipitation (MAP) in determining amphibian richness in simple regression models corresponds to findings of several regional analyses (Schall & Pianka 1977, 1978; Seymour et al. 2001; Hawkins et al. 2003; Qian et al. 2007; Ortiz-Yusty, Páez & Zapata 2013). However, it deviate from those of several others studies that found ambient energy (Allen, Brown & Foody 2002), water-temperature dynamics (Qian 2010) and ecosystem productivity (Buckley & Jetz 2007) to correlate best with amphibian richness. In a meta-analysis, Hawkins et al. (2003) found water and the availability thereof to be the primary determinants of amphibian distributions across tropical realms (i.e. Neotropics) while the energy availability restricts diversity across temperate realms (Palearctic, Nearctic). The results presented here add further support to these conclusions which are consistent with the physiology of anurans: in tropical realms where energy is more sufficient, water availability will determine the number of species co-occurring owing to reproductive (to lay eggs) and survival (to avoid desiccation) needs. The implication of this is that anthropogenic activity and climatic events which adversely influence ecosystem hydrology will undoubtedly prove detrimental to anurans. From a conservation perspective, this is important to note because such areas will need to be prioritised if current anuran biodiversity is to be preserved for future generations.

Furthermore, the significance and amount of variation explained by the multiple regression models also suggest that multiple environmental factors act in concert to constrain anuran alpha diversity. In particular, the substantive variance explained by the local regression model (92%) suggests that present day anuran distribution patterns are closely in tune with current environmental factors. However, the unexplained variation accounted for by both the global (38%) and local model (8%) may in part reflect environmental, historical and ecological factors not directly tested here. These findings corroborate those of Buckley & Jetz (2007) who found the influence of environmental processes at a global scale, irrespective of biogeographical realm, to surpass those of historical and ecological factors in determining anuran richness. From a conservation management perspective, the tight coupling of anuran distribution patterns to current environmental conditions is ominous because it implies that environmental change, induced by either climate change or anthropogenic influence, will undoubtedly accelerate anuran declines.

3.5.5 **Determinants of beta diversity**

3.5.5.1 OLS

With the OLS model, significant relationships ($p < 0.05$) between beta diversity and environmental correlates were obtained (Table 3.9). However, because of the significant Koenker (BP) scores associated with the OLS model, the robust probability associated with each environmental variable were inspected to determine whether Pearson probabilities were valid. Based on the significant robust probabilities ($p < 0.05$), all modelled variables presented a significant contribution in accounting for the variation in the spatial turnover of anurans across extent.

All environmental variables displayed a negative relationship with beta diversity except for topographic complexity (TOPOSD) which was positive. MAT presented the steepest coefficient slope while NPP presented the gentlest. The positive relationship between TOPOSD and beta diversity implies an increase in biotic dissimilarity with increasing topographic complexity. Similar behaviour between topographical complexity and amphibian turnover has recently been documented by Baselga et al. (2012). The slope of the model intercept ($\beta_0 = 1.44$) was also positive which indicates an increase in biotic dissimilarity towards environmental extremes.

Table 3.9: OLS model parameters for beta diversity.

	Diagnostic			
	Coefficient	**Robust t-value**	**p - value**	**Robust p - value**
Intercept	1.44	10.20	< 0.05	< 0.05
MAP	-0.08	-10.33	< 0.05	< 0.05
MAT	-0.41	-13.35	< 0.05	< 0.05
TRange	-0.17	-18.50	< 0.05	< 0.05
NPP	0.03	-3.77	< 0.05	< 0.05
TOPOSD	0.05	6.35	< 0.05	< 0.05
Model Performance		**z-score**		
Adj. R^2	0.12			
AICc	1683.78			
Morans I	0.67	72.76	< 0.05	
Koenker	292.99		< 0.05	
Jacque Bera	318.44		< 0.05	

In terms of model fit, the OLS indicates that environmental variables could explain only 12% of the variation in beta diversity (adj. R^2 = 0.12). This was expected because the ratio of low beta diversity exceeds that of high beta diversity. That is, high beta diversity comprises only a small portion of the study extent in relation to low beta diversity. The poor performance of global linear models in explaining beta diversity has been noted by Di Virgilio et al. (2014). Based on the explanatory regression, TOPOSD was the single strongest predictor explaining 5% of the variation in anuran beta diversity while the remaining predictors all explained less than 1% in the deviance of beta diversity across the study extent.

Furthermore, the Jacque-Bera statistic, which checks model accuracy by testing normality of the residuals, was significant ($p < 0.05$) indicating the presence of model bias. To source the bias, the standardised residual map (Figure 3.17 & Figure 3.18), residual scatter plot (predicted to observed model values) (Figure 3.19) and residual histograms produced by the model were inspected.

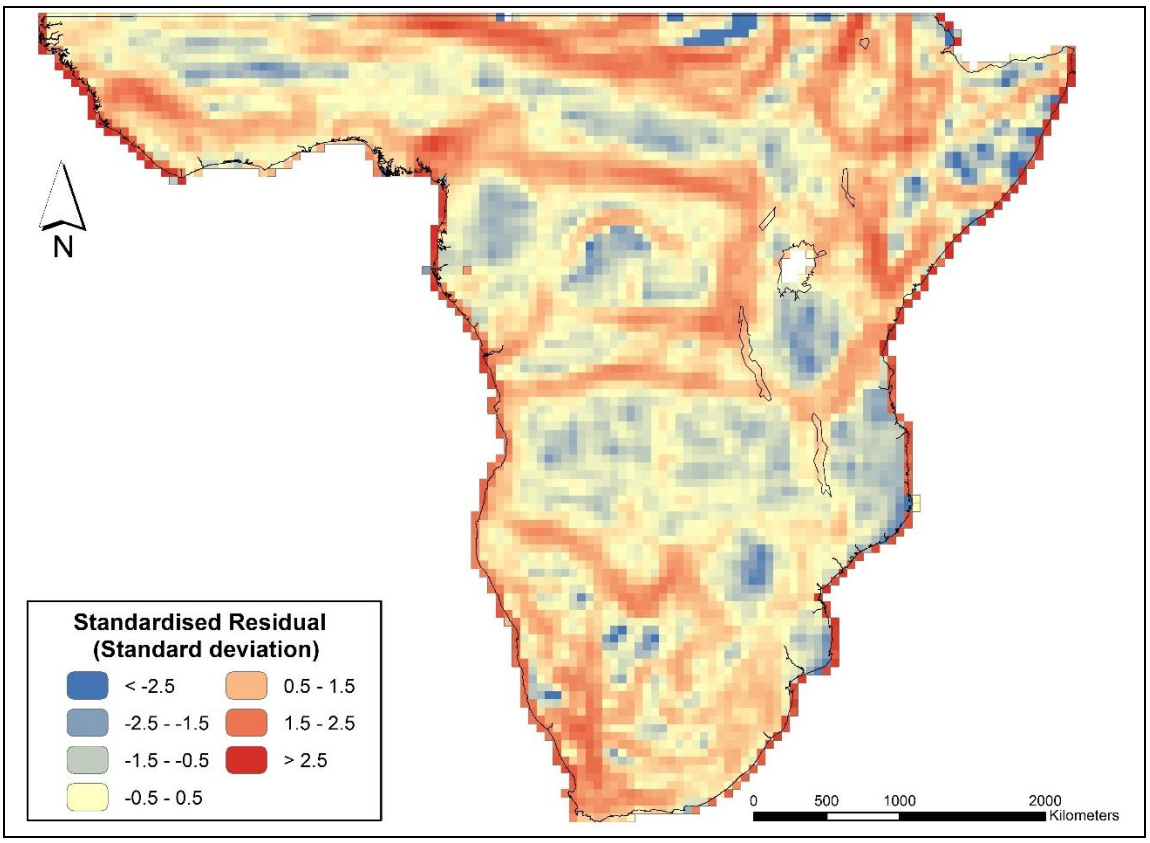


Figure 3.17: Spatial distribution of standardised residuals produced by the OLS model for beta diversity across the study.

The standardised residuals of the OLS regression showed that model consistently over-predicted in areas of high diversity (i.e. transition boundaries) and under-predicted in areas of low diversity (interior of biogeographic provinces) (Figure 3.18).

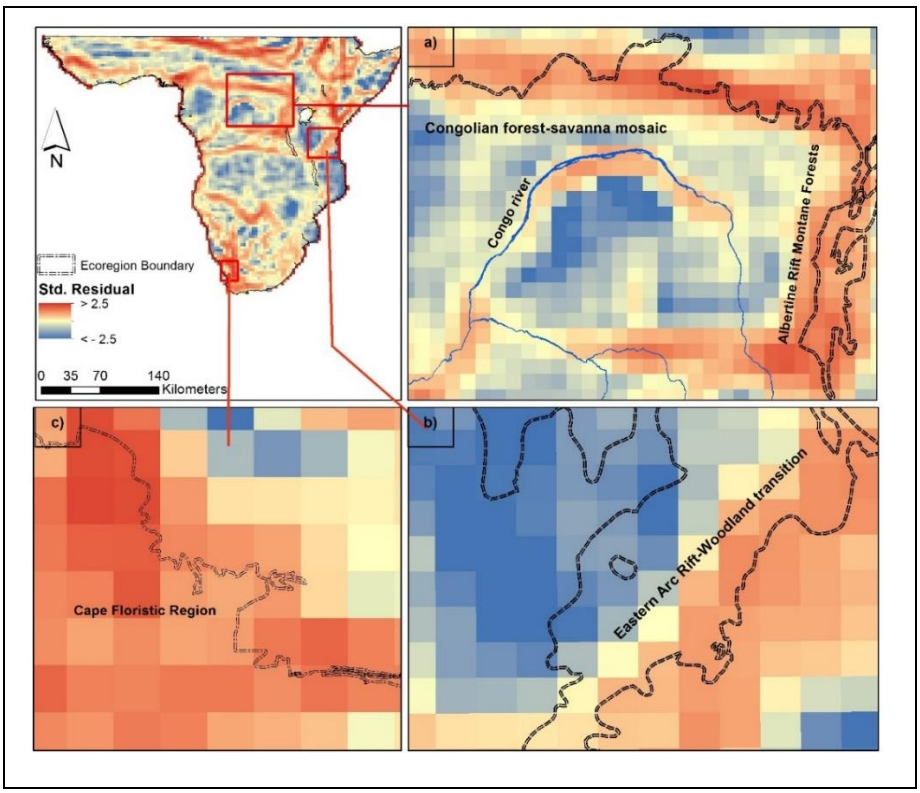


Figure 3.18: The OLS model over predicted along biotic transitions in the (a) Congo basin (b) Eastern Arc Mountains and (c) Cape Floristic region.

Considering the significant high Morans I score associated with the OLS residuals (*i* = 0.67, *z*-score = 72.76, $p < 0.05$), model bias was most likely due to misspecification, specifically omitted or missing predictor variables.

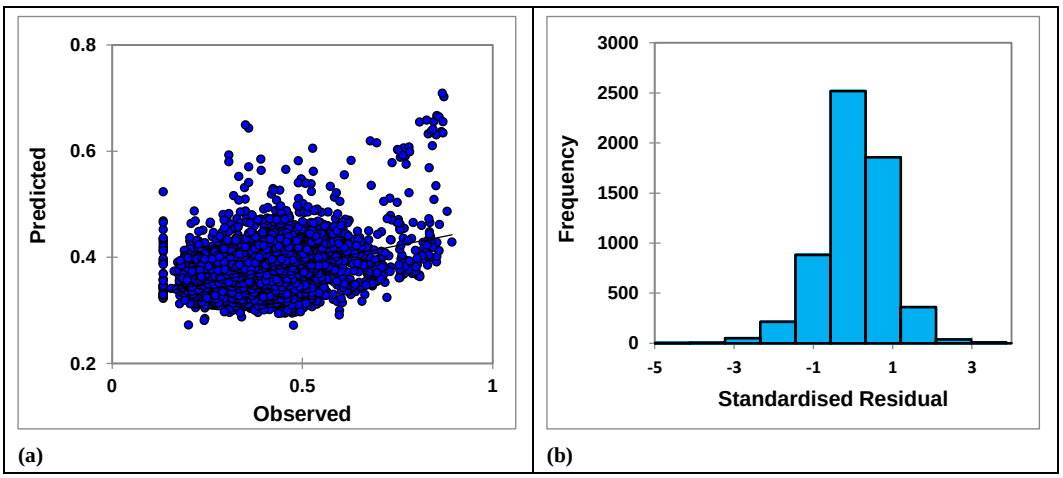


Figure 3.19: Scatter plot of (a) predicted to observed model values for beta diversity with corresponding (b) standardised residual plot for the GWR model.

3.5.5.2 GWR

For the GWR, significant relationships between beta diversity and environmental correlates were obtained ($p < 0.05$). All environmental correlates presented significant contribution towards the model. The strength and direction of the median coefficient slopes obtained in the GWR model were similar to those of the OLS. Specifically, the coefficient slopes of the predictors were all negative expect for TOPOSD which was positive. However, in contrast to the OLS model, the

median intercept of the model presented a negative fit. This relationship however may be biased as the condition numbers, a measure of local collinearity, associated with the model intercepts frequently exceeded the unstable parameter threshold (CN > 30). In the presence of collinearity, coefficients produced by the model may be biased or unstable.

Table 3.10: Median coefficient slopes of predictors in GWR model.

	MAP	MAT	TRange	NPP	TOPOSD
Median	-.0001311	-.01425	-.15227	-.000000484	**.01158174**

Model parameters for the GWR were distributed variably across the study extent (Figure 3.20). Simple interpretation of the spatial distribution of predictor coefficients was not straight forward except for TOPOSD. In general, coefficient slopes of topographic complexity (TOPOSD) were steep and positive in topographically complex areas such as the Ethiopian highlands, East African Rift Valley and Cape Floral Kingdom east through to the Drakensberg mountain range of South Africa. Gentle to steep negative slopes were observed across topographically stable areas such as the Miombo Plateau and Sahel region to the north.

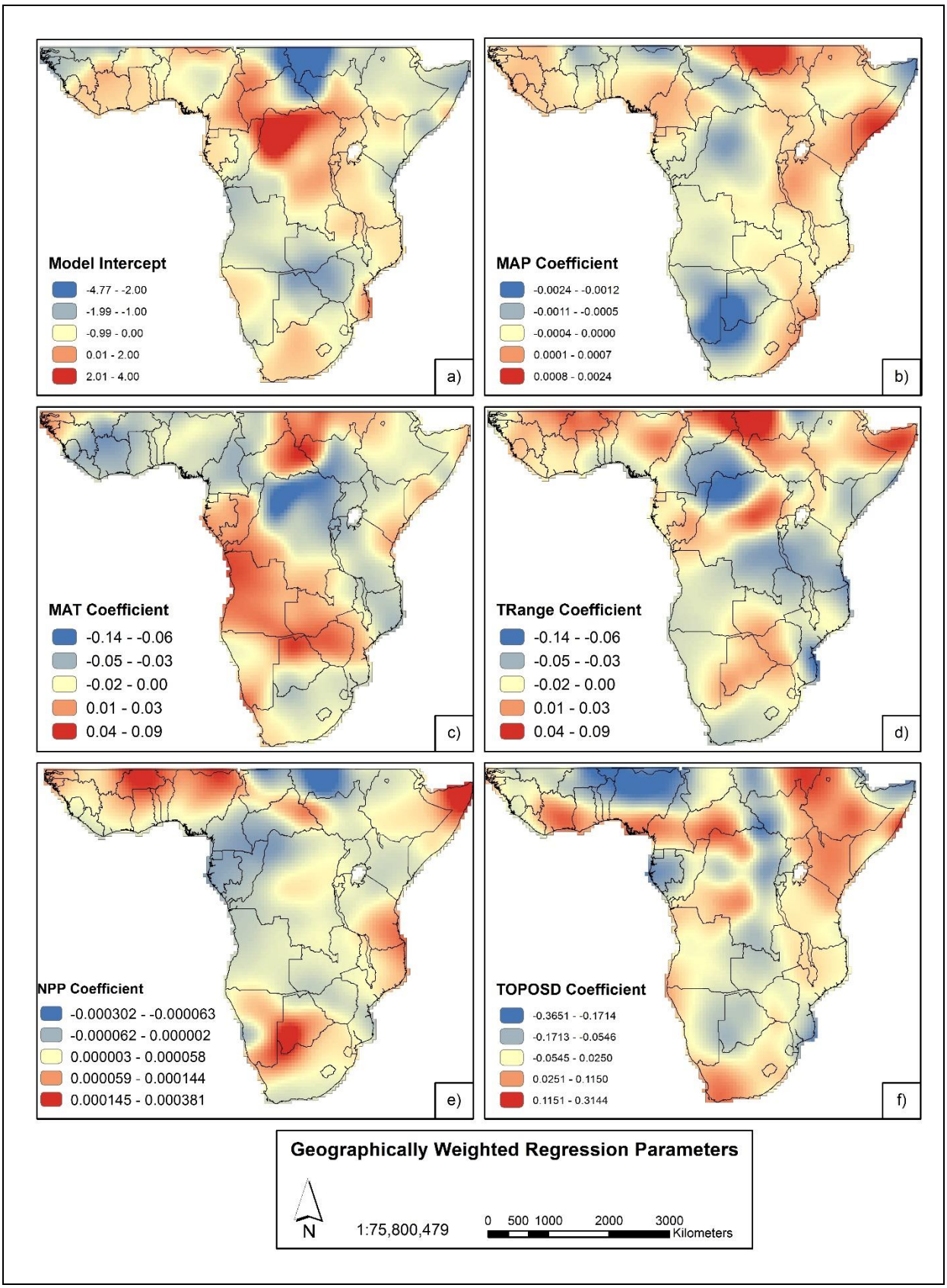


Figure 3.20: Spatial variation in (a) model and (b – f) predictor coefficients numbers across the study extent at 0.5º × 0.5º resolution.

Overall, the GWR regression could explain 54% of the variation in beta diversity across the study extent, a substantial improvement over the 12% explained by the OLS model. The spatial distribution of the local r^2 (Figure 3.21a) however showed that model fit was spatially variable. Again, direct interpretation was precluded but high model fit was observed in west, central and

southern portions of the study extent. Most of East Africa had poor model fit indicating missing explanatory variables, perhaps related to the complex geologic and climatic history of this region. For example, Murphy et al. (2015) showed that beta diversity of a region may correlate highly with the geological history of that region. Similarly, Baselga, Gomez-Rodriguez & Lobo (2012) linked global amphibian turnover to paleoclimate.

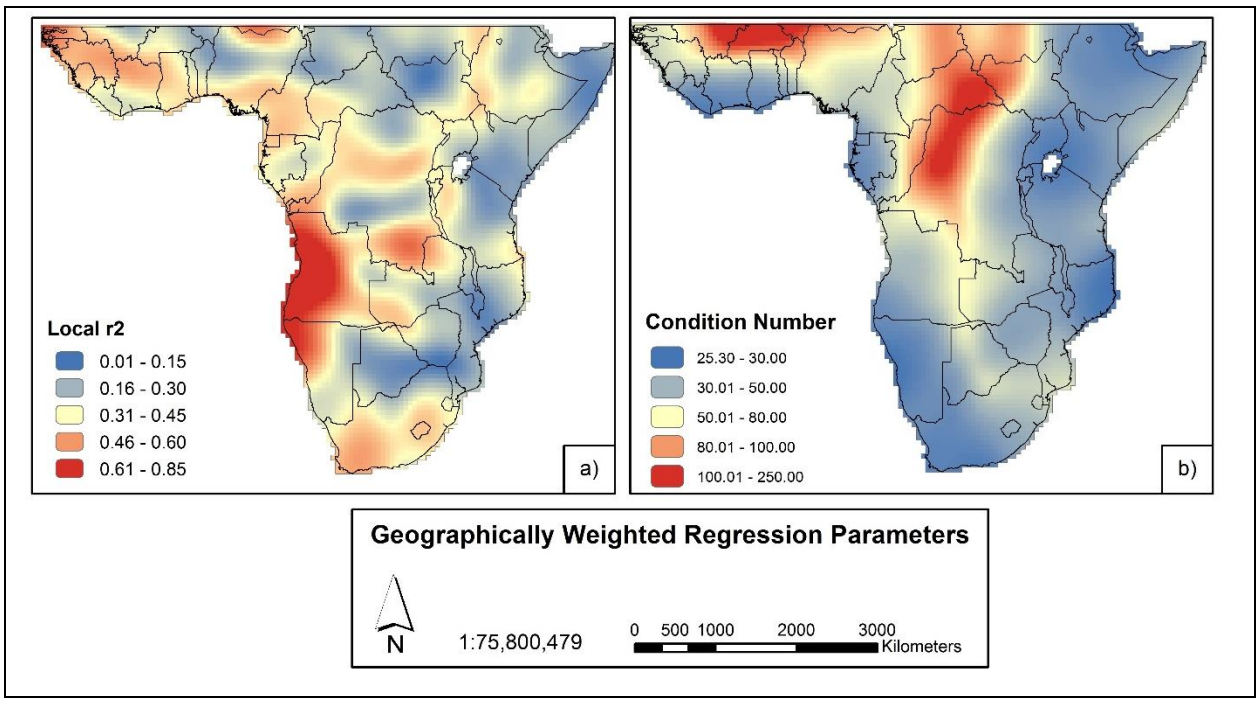


Figure 3.21: Spatial distribution of the (a) local r^2 and (b) condition number across the study extent.

Furthermore, the lower Morans I value associated with the GWR showed that this model accounted better for spatial dependence amongst the residuals than the OLS model. However, inspecting the standardised residual map (Figure 3.22) of the GWR showed that like the OLS, though to a lesser degree, the GWR model consistently overestimated in areas of high species turnover (i.e. along biogeographical boundaries) and underestimated in areas of low species turnover (i.e. within biogeographical provinces).

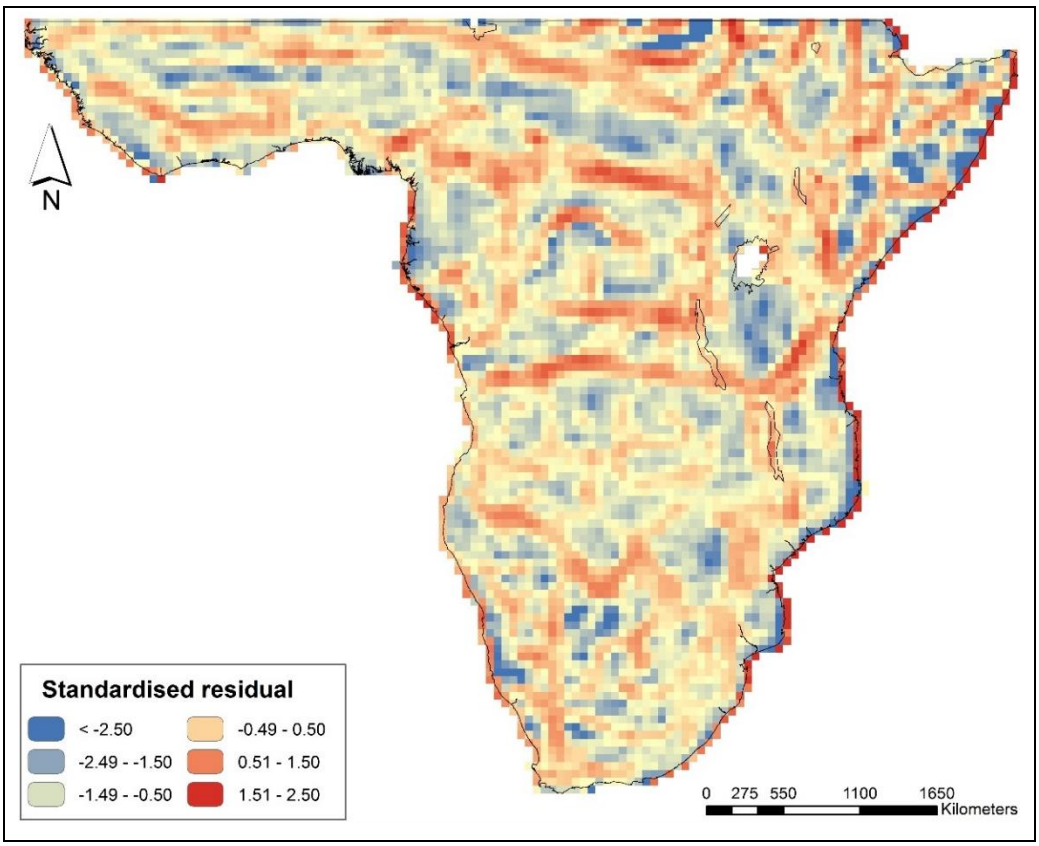


Figure 3.22: Spatial distribution of standardised residuals produced by the GWR model across the study extent. Shades of red indicates where the model overestimates while shades of blue indicate where the model underestimates.

This significant autocorrelation suggests model misspecification by missing predictor variables. This affirms the conjecture of missing explanatory variables in East Africa where low model fit was obtained. A scatter plot (Figure 3.23a), depicting predicted to observed model values, demonstrates the improved model fit presented by the GWR while the histogram of standardised residuals (Figure 3.23b) shows the reduced bias in the models residual structure in. Specifically, a greater proportion of the residuals fall within ±2.5 standard deviations from the mean than those of the OLS model.

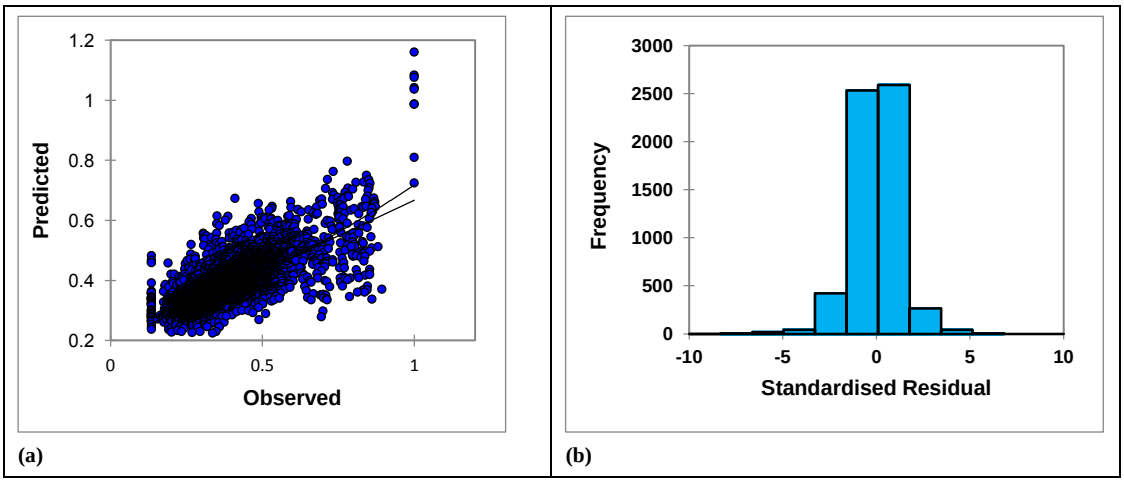


Figure 3.23: Scatter plot of (a) predicted to observed GWR model values for beta diversity with corresponding (b) standardised residual plot. Values for the scatter plot have been back transformed to facilitate interpretation.

3.5.5.3 Model Comparison: Overview

When comparing the GWR model to the OLS model (Table 3.11), the GWR recorded the lowest AIC score (78.12), which indicates that this model had a better balance between the amount of variance explained and the number of variables included than the OLS model (AIC = 1801.88). The GWR also corrected best for spatial dependence in model residuals as indicated by the significantly lower Moran's I scores (GWR = 0.67, OLS = 0.92). The improved performance of the GWR in accounting for spatial autocorrelation has been attributed to its ability to allow for spatially variable relationships and the consideration of neighbour influence during model calibration (Propastin et al. 2008; Sheehan, Strager & Welsh 2013).

Table 3.11: Comparison of the OLS to GWR performance using the set evaluation criteria.

Model	Adj. R^2	AIC	Morans I	z-score	P value
OLS	0,12	1683.78	0.67	72.76	$p < 0.05$
GWR	0,54	78.12	0.52	56.96	$p < 0.05$

Furthermore, the GWR also explained the most variation in spatial turnover, specifically the GWR could explain 54% to the 12% accounted for by the OLS. When considering the model evaluation criteria, the GWR consistently outperformed the OLS model in all criteria further showing the data were better suited to a local and not a global model. The improved performance of the local GWR to global OLS documented here are concordant with several other published studies (Fischer & Getis 2010; Sheehan, Strager & Welsh 2013).

3.5.6 **Comparison with respect to other studies**

Results from the regression models suggest that environmental variables are significantly correlated to and can explain a considerable amount of variation in beta diversity. In particular, the local GWR model suggests that environmental variables are strong local determinants of beta diversity. These findings correspond to several parallel analyses and asserts further empirical support to the environment-beta diversity which suggest a decrease in biotic similarity with increasing environmental dissimilarity (Legendre et al. 2009; Chen et al. 2011; Wang et al. 2012; Fitzpatrick et al. 2013; Podani & Schmera 2016). The rationale here is that greater environmental differences between areas result in greater variation in environmental niches between those areas, thus leading to higher rates of beta diversity (Davidar, Puyravaud & Leigh 2005).

As the methods used to quantify beta diversity with environmental processes, as well as the environmental variables used to measure environmental dissimilarity vary among studies, results from different studies for different taxa documented across different spatial scales are difficult to compare. Buckley & Jetz (2008) found amphibian turnover increased with increasing environmental distance but noted that high beta diversity could occur in the absence of high

environmental dissimilarity. This suggests that the contribution of neutral processes in the structuring of beta diversity patterns need to be considered if the true mechanisms underlying the structuring of beta diversity are to be understood. Mcdonald et al. (2005) support this notion finding spatial turnover in North American mammals to occur across homogenous climatic zones.

In the regression models, TOPOSD was the single strongest predictor of beta diversity. The predictive power of topographic variability concurs the findings of Baselga, Gomze-Rodriguez & Lobo (2012) who found topographic variability as a primary determinant of amphibian beta diversity south of 37˚N latitude. The influence topographic complexity exerts on the structuring of regional beta diversity patterns may be attributed to three factors. First, amphibians possess limited dispersal capacity and thus the increase in compositional dissimilarity with increasing topographic complexity may reflect the discrete ecological response of anurans to dispersal barriers. Second, at a mesoscale, topographic complexity is associated with increased habitat heterogeneity presented over short spatial distances. This may promote beta diversity through niche filtering. Third, in the past, topographically complex areas have experienced more stable climatic conditions than their surrounding areas serving as refuges to species promoting the persistence of some species (niche conservatism) and the adaptation (niche adaptations) and evolution (speciation) of others.

Furthermore, the discrepancies in the results obtained by the global OLS and local GWR highlight the importance of the consideration of statistical techniques in determining pattern-process relations. In particular, the results presented here suggest that global models are useful to determine whether there are relationships between beta diversity and environmental variables while local models may reveal more detail about locally associated processes which shape patterns of beta diversity.

In conclusion, the significant relationships between beta diversity and environmental variables suggest that species turnover of anurans across space is, to some degree, spatially structured. From a conservation perspective, this spatial structure is important because it implies that species distributions and co-occurrence (α-diversity) patterns are non-random in nature and in order for nature to recreate and maintain itself, this spatial organisation needs to be preserved. Proper ecosystem management can thus only be implemented if this spatial organisation is conscientiously considered and evaluated when conservation strategies are drawn up.

3.6 CONCLUSION

In summary, the present study examined the influence of the contemporary environment on patterns of alpha and beta diversity. Results obtained from multivariate spatial and aspatial regression models found environmental variables to be strong determinants of diversity. For alpha diversity, precipitation was the strongest predictor of anuran richness patterns, reaffirming the crucial importance of water availability for the taxa. The importance of water reflects the physiological limits it imposes on anurans that require it for both survival and reproduction. With their acute sensitivity to environmental conditions, anurans may thus be at greater risk of extinction than other vertebrates if dramatic climatic shifts result in significant precipitation anomalies or ecosystem hydrology across the Afrotropics is disrupted.

For beta diversity, topographic variation accounts for most variation in biotic heterogeneity. Anurans with limited dispersal capacity, particularly narrow ranged ones will be at greater risk of population declines or extinction because they cannot readily disperse to climatically favourable sites. In addition, results from two regression models show that it is important to analyse data using different techniques as different spatial structures and relationships may emerge with the consideration and implementation of different spatial techniques.

CHAPTER 4: THE SCALE RESONANCE OF SPECIES RICHNESS AND TURNOVER

4.1 INTRODUCTION

Spatial patterns and spatial scale are inseparable concepts in ecology and biogeography. This is because different patterns manifest at different scales owing to different processes operating across different spatial and temporal scales (O'Neill et al. 1991). That is, patterns and pattern-process relationships observed at one spatial scale may thus not be actual, but instead be an artefact of the scale of analysis.

This comprises the focus of the current chapter. Specifically, it looks at the effect of changing spatial scale on patterns and environmental drivers of anuran diversity. Here, the methodology and analysis presented in the previous chapter are repeated but across multiple spatial scales. Essentially, this component entails the validation of the results obtained in the previous chapter when considering multiple spatial scales. As this chapter is linked to the previous, repetition may be present.

Three contrasting relationships have been proposed for the spatial scaling of species diversity (Keil et al. 2012). Type I relationship suggests that species richness increases linearly as function of the area sampled. Type I relationship is related to the slope (***z***) of species-area relationship which increases with increasing area. In return, species turnover is expected to also increase with increasing grain because a larger value of ***z*** corresponds to a faster accumulation of species (i.e. inclusion of rare species). Type II relationship suggests that species turnover is expected to decrease with increasing grain because environmental differences are smaller at broader spatial grains due to a decrease in within sample variance owing to spatial averaging (Lennon et al. 2001). Environmental differences are expected to explain a greater degree of variation in species turnover at broader spatial scales. Type II relationship is nested in niche theory (Keil et al. 2012). In Type III relationship, species turnover is expected to increase with increasing spatial grain because the spatial distance (lag) between sampling units are greater i.e. distance-decay relationships (Nekola & White 1999; Mcdonald et al. 2005; Qian & Shimono 2012). This relationship is taxa specific and is expected to be higher for taxa with lower dispersal capacity such as amphibians.

Although there is an agreement that beta diversity determines the relationship between local and regional species richness (Bishop et al. 2015), there exists no overall consensus of the nature of the spatial scaling of beta diversity. This is because beta diversity is dependent on an array of factors including the taxa of interest, geography of the sampling extent and geographic gradients

persisting across the sampling extent. This research component tests which type of relationship(s) is at work for amphibians of the Afrotropics.

4.2 METHODS

4.2.1 SCALING

As the spatial resolution of data increases so does its variability and complexity (Wiens 1989). As grain size increases, so does the area it covers on the ground and vice versa. For example, at fine scale, a landscape can comprise discernible individual trees, but at a coarser scale only broad classes of vegetation can be distinguished. One suggested approach to addressing this issue is to aggregate data systematically and then examine if and how patterns change across scales statistically (Francis & Klopatek 2000).

4.2.1.1 Sampling design

First, a hierarchical sampling framework was adopted to examine the effect of changing scale, in particular spatial grain, on patterns and drivers of anuran diversity. This framework is based on hierarchy theory which postulates that ecosystems do not exist in isolation but as nested hierarchies, with each ecosystem nested within another, from smaller to larger (Gibson & Ostrom 1998; Wu 2004).

The hierarchical framework comprised the use of a hierarchical pooling scheme using three sets of equal area quadrats of sizes 0.5° × 0.5° (56 × 56 km), 1° × 1° (112 × 112 km) and 2° × 2° (224 × 224 km) resolution near the equator using an equal area projection[1]. Hierarchical pooling entails the progessive aggregation of sampling units from smaller to larger ones such that each successive set is a complete subset of the previous (Figure 4.1). During the analysis, only grain size was altered while extent remained fixed.

[1]Initially, four-grain sizes were used 0.25°, 0.5°, 1°, 2°. However, the 0.25° grid set was ommitted from the study owing to computational issues associated with the regression models. This grain however was used to address the second obective set for the chapter.

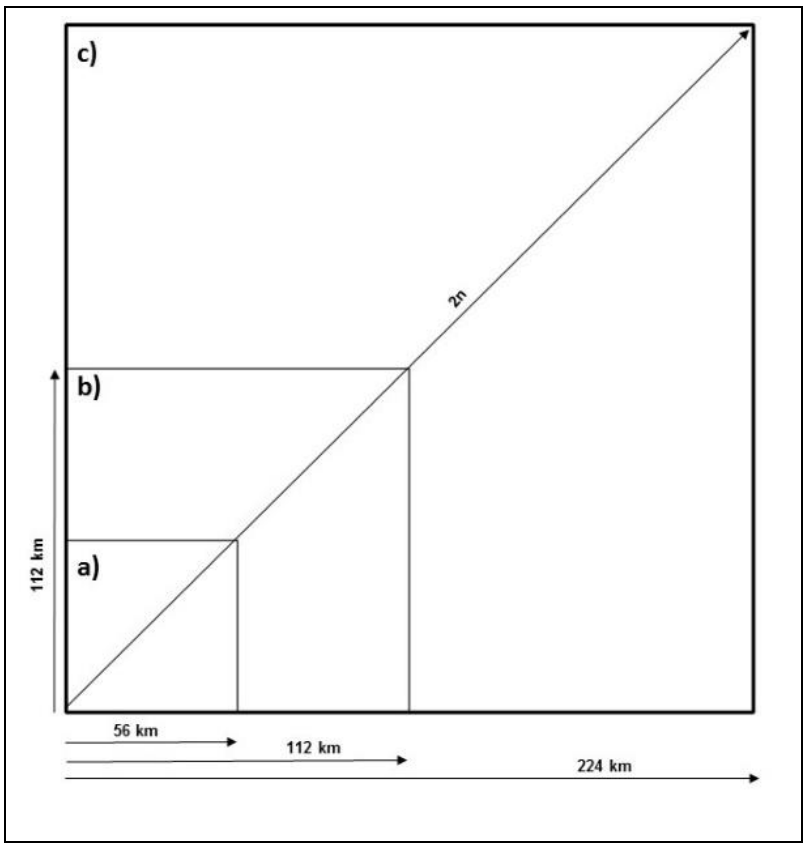


Figure 4.1: Hierarchal sampling scheme modelled after Arita & Rodriguez (2002) used to examine the effect of grain on diversity. Quadrats were increased such that quadrat **a** was a subset of **b** and (**a – b**) were both subsets of the largest quadrat **c**.

4.2.2 **Patterns of diversity**

Next, alpha and beta diversity for the grids were established using the methods decribed through sections 3.1.2 - 3.1.4 of the research document. Briefly, this included: establishing the presence/absence of species in a quadrat by superimposing and intersecting the species dataset with the grid systems. This was then followed by creating a moving window to identify and create spatial neighbourhoods assocaited with each focal cell to derive diversity. Lastly, zeta diveristy paritioning was used to quantify diversity. In particular, species richness (α-diveristy) was quantified using the first order zeta component while turnover (β-diversity) was quantified using the Sorensen dissimilarity index (β_{sor} = 1 – Sorensen) derived from zeta diversity components as follows ($\zeta_{sor} = 1 - \zeta_2/\zeta_1$).

Maps depicting the spatial distribution of the two diversity components were then created using ArcGIS to observe and compare the effect of changing spatial grain on diveristy patterns.

4.2.3 **Summary statistics – Underlying distributions**

Parviainen, Luoto & Heikkinen (2010) suggest that summary statistics (SS) are suitable empirical units for quantifying changes in ecological data across spatial scales. For this reason, summary statistics (SS) were used to assess how changing spatial grain influenced the underlying distribution of diversity data (i.e. empirically). The SS used included the minimum (min), maximum (max), mean and standard deviation (SD) of α-diversity (= γ-diversity) and

turnover (β-diversity). Values associated with coastal quadrats were removed from analysis because of the area based bias they may induce (Whittaker, Willis & Field 2001).

Histograms for the four SS were then created to aid in visual interpretation of the scaling results. As biotic patterns result from both linear and nonlinear interactions, trend lines were fit through the histogram points to observe the parametric behavioural pattern of the data across spatial grain. SS and their respective histograms were derived using a trail version of Microsoft XLSTAT v4.0.

For both alpha and beta diversity, all metrics were expected to increase as a function of spatial grain due to the species-area relationship ("more area permits for more species") (Lark 2011). No discrimination was made to the parametric nature of the increase. An exception however, was the standard deviation metric which was expected to decline as a linear function of spatial grain owing to a reduction in within sample variance associated with larger spatial grains (Marceau & Hay 1999; Whittaker, Willis & Field 2001).

4.2.4 **Pattern-process relationships across grain**

Pattern-process relationships are scale dependent. That is, as the scale of the analysis is altered so is the relationship between pattern and process. Pattern-process relationships observed at one scale may thus not hold at another (Field et al. 2008). Ecologists often conduct studies at, and draw conclusions from, a single scale of analysis. However, as noted by Field et al. (2008), such deductions may prove misleading or even erroneous. To truly understand and disentangle mechanisms multiscale studies are needed.

Thus, to examine how the relationship between patterns of species diversity (richness and turnover) and environmental processes vary with varying spatial grain, the regression analysis described in Section 3.4 entitled "Modelling" were repeated for the larger spatial grains.

As a priori, the pre-processing steps mentioned in section 3.1.6, taken for the environmental variable were repeated for the coarser grains. In chronological order, this included: the scaling of the environmental correlates to the size of the coarse grid sets using the resampling and aggregation algorithms previously described; the extraction of the aggregated environmental data from these layers using centroids assigned to quadrats in a Euclidean space; and the application of statistical treatments to the data (i.e. transformations). These analyses were only conducted for the same environmental variables selected in section 3.2 as determined through correlation analysis and model selection procedures using the $0.5^{\circ} \times 0.5^{\circ}$ grid. This was because this grid set presented the finest spatial resolution of the study. The environmental variables included in the

regressions across grain were MAP, TRange and NPP for alpha diversity and MAT, MAT, TRange, NPP, TOPOSD for beta diversity.

Again, alpha and beta diversity were the response variables in the regressions while the environmental variables were the predictors. Both an Ordinary Least Squares (OLS) and Geographically Weighted Regressions (GWR) were performed on the data. For the GWR, parameters were the same as before (i.e. fixed spatial kernel, 50 neighbours as bandwidth parameter). Different bandwidths were experimented with but the best results remained relatively consistent for the 50 neighbour bandwidth and thus only results obtained at this bandwidth were retained. To evaluate the regression models (OLS & GWR) for both alpha and beta diversity, the same model evaluation criteria set in section 3.4.33.4.3 of the research document was applied to the larger spatial grains. This included the adjusted R^2 to evaluate model fit, the AIC score to evaluate model performance and the Moran's I parameter to evaluate the spatial autocorrelation of the data.

The final criteria used to check for the effect of changing spatial grain on pattern-process relations included: (1) the coefficient slopes of the regression models as well as the predictors within the regression models, (2) the significance of each predictor within the model and the model itself, (3) and the three set model evaluation criteria.

4.3 RESULTS: PATTERNS OF DIVERSITY

4.3.1 Alpha diversity

Figure 4.2 shows the spatial distribution of alpha diversity for the study extent, derived using the first order zeta component (ζ_1) across the three spatial scales. Dark red indicates areas of high species richness while dark green indicates areas of low species richness. The hue of each colour corresponds to the relative intensity of the number of species (i.e. light green has higher species richness than dark green).

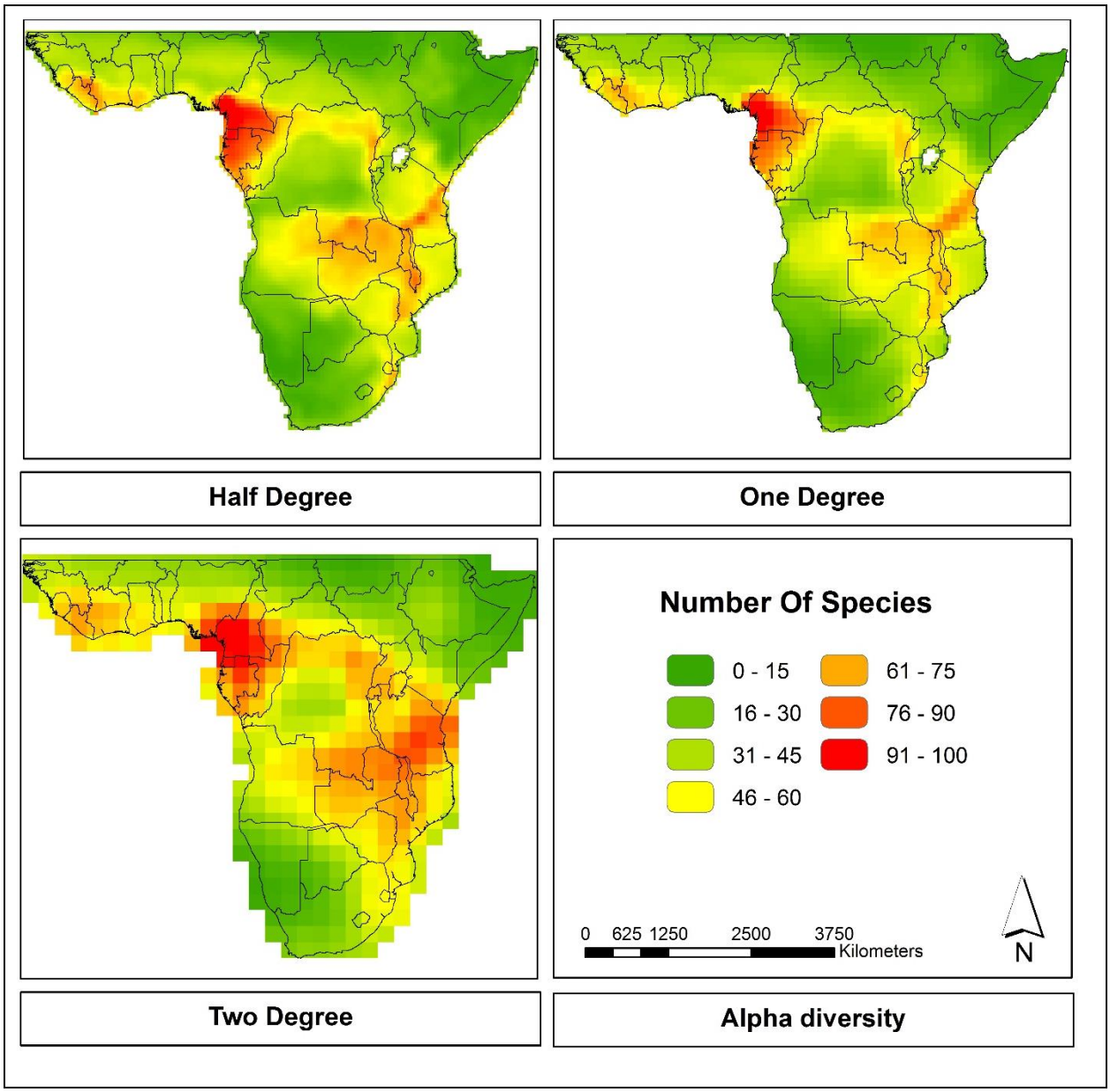


Figure 4.2: Spatial variation in alpha diversity across the study extent at (a) 0.5˚ × 0.5˚ (b) 1˚ × 1˚ and (c) 2˚ × 2˚ resolution.

For all grain sizes, tropical Africa exhibited the highest levels of alpha diversity. In particular, the Congo Basin and Gulf of Guinea from southern Gabon through to the Niger Delta have high species richness. The Cameroonian Highland Forests remained the most species rich area (± 100 species). High levels of alpha diversity were also observed in West Africa near the Guinean

Savanna-Forest complex from northern Senegal south through eastern Liberia in the vicinity of Mount Nimba. In East Africa, the Eastern Arc Mountains of Tanzania extending from the Pare Mts. in the north, south towards the Udzungwa Mts. near the Malawi also demonstrated high alpha diversity.

Intermediate alpha diversity (34 - 70 species) was observed across much of South-eastern Africa. In general, alpha diversity was higher in moist regions and attenuated towards latitudinal extremes with the arid Sahel and Namib – Kalahari Desert regions recording the lowest number of species (< 15 species).

Broadly, the overarching patterns of alpha diversity remained relatively consistent across grain. However, as the spatial grain of the sampling set was increased, the homogeneity of the spatial patterns also increased. That is, as smoothing increased the spatial heterogeneity and variance of the system was decreased. This was particularly true for fine scale patterns observed at the finer half-degree grain which appeared visibly smoothed out at the coarser two-degree grain. Extracting a sample of size equal to the two-degree quadrat from the study area for all three grain sizes shows how the variance of system is decreased as grain size is increased (Figure 4.3).

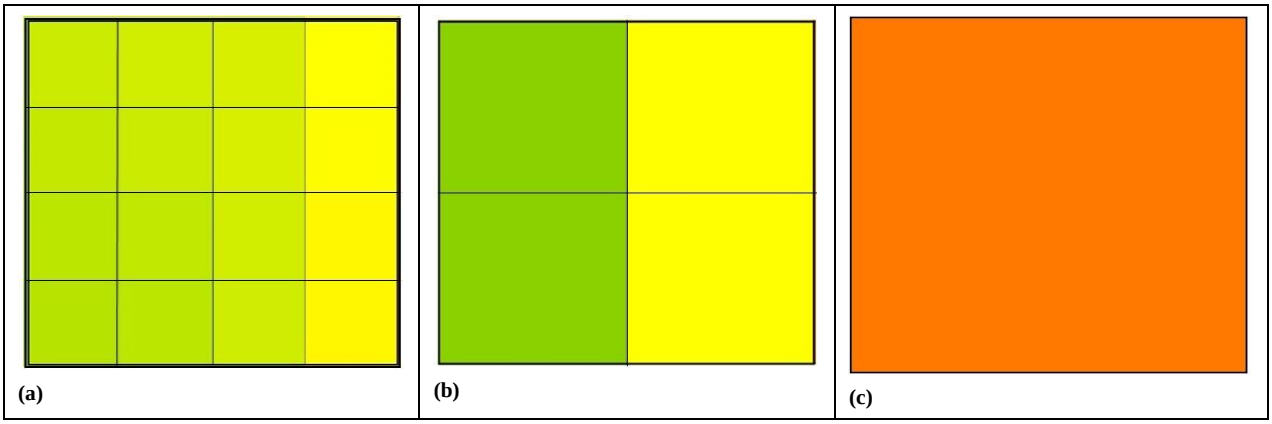


Figure 4.3: Spatial heterogeneity of the system is visually reduced from (a) 0.5˚ × 0.5˚ (b) 1˚ × 1˚ and (c) 2˚ × 2˚ resolution.

4.3.2 **Beta diversity**

Figure 4.4 shows the spatial distribution of beta diversity across the study extent as measured by the Sorensen index derived per zeta diversity partitioning. Dark red indicates areas of high biotic dissimilarity and dark green indicates areas of low biotic dissimilarity. The hue of each colour corresponds to the intensity of biotic dissimilarity i.e. yellow has lower biotic dissimilarity than orange.

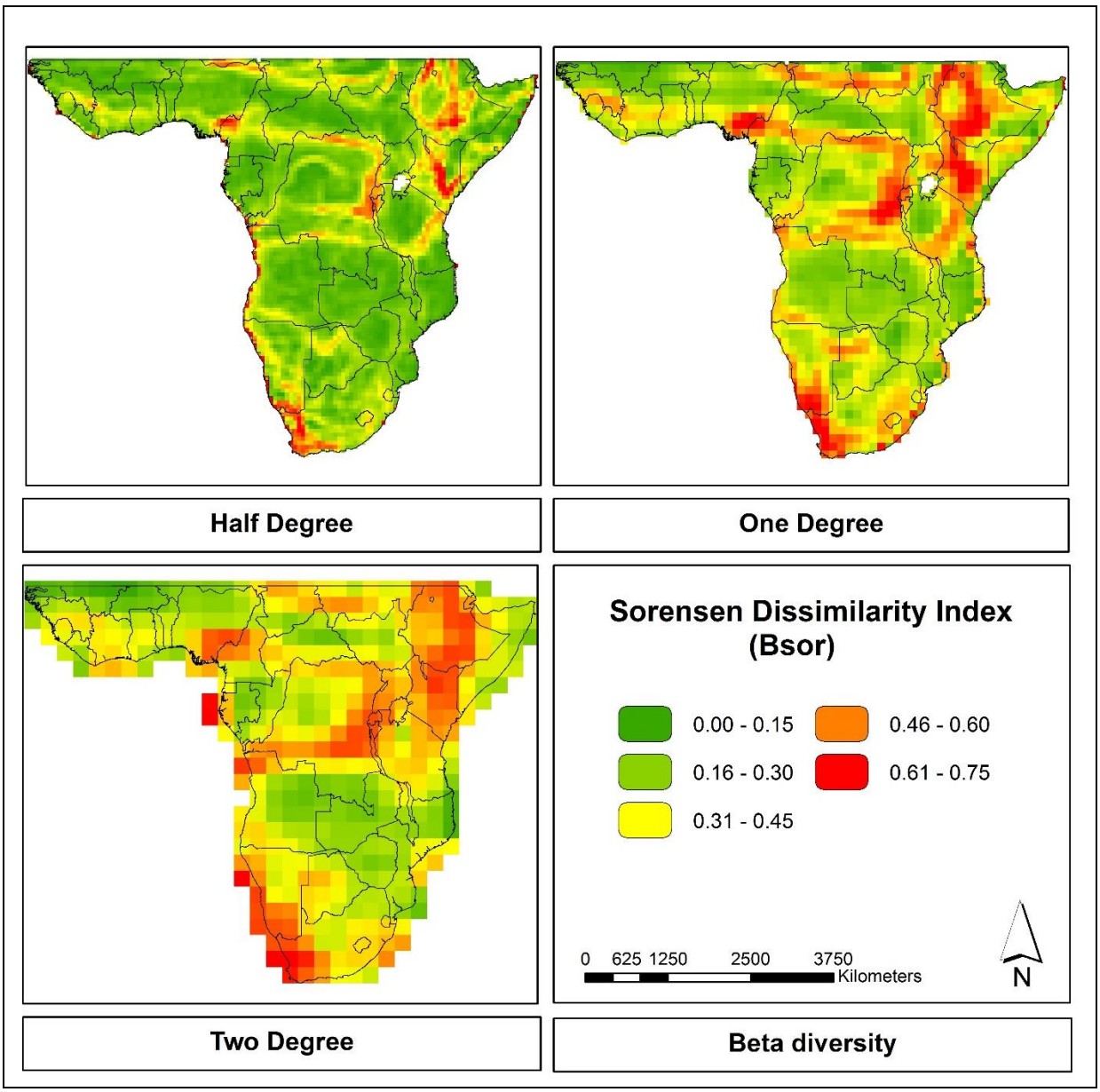


Figure 4.4: Spatial variation in beta diversity across the study extent at (a) half (b) one and (c) two degree resolution. As with alpha diversity, patterns of turnover broadly remained consistent across grain.

In contrast to patterns of species richness, where diversity is concentrated in tropics, high levels of beta diversity occurred across a range of latitudes. East Africa, from the Ethiopian highlands through the Congolian – Albertine montane forest complex, were areas of high beta diversity. High beta diversity was also observed in north central Africa distributed in a front like manner along the Northern Congolian Forest-Savanna Mosaic border west through the Cameroonian highlands to the Guinean montane forest complex (Figure 3.11). In central Africa, specifically the Congo Basin, regions flanking the Congo river also presented high beta diversity (β_{sor} = ± 0.4).

The Miombo woodlands/Congo forest transition presented intermediate beta diversity. In contrast, beta diversity was generally low across much of Southern Africa, with the exception of the Cape Floristic region north through the Karoo region where high beta diversity, equivalent to

the maximum, was recorded ($\beta_{sor} = 0.7$). Low beta diversity ($\beta_{sor} < 0.1$) was also observed across much of the northern Afrotropics along the Sahel region and Southeastern Africa across the Zambezi plateau. Overall many areas of high beta diversity coincided with known biogeographical transitions, while low beta diversity persisted within biogeographical provinces.

Similar to that of alpha diversity, the overarching patterns of beta diversity remained relatively consistent across grain. However, as the spatial grain was increased the spatial homogeneity of the spatial patterns also increased. Spatial variance was thus lower at the coarse two-degree grain than the finer scale half degree grain.

4.4 GRAIN DEPENDENCY: SUMMARY STATISTICS

4.4.1 Alpha diversity

Table 4.1 and Figure 4.5 show the variation in the summary statistics of alpha diversity across spatial grain.

Table 4.1: Summary statistics extracted for alpha diversity across spatial grain.

Statistic	Grain Size (degrees)			
	Quarter	Half	One	Two
Number of observations	23152	5943	1501	417
Minimum	3	3.7	4	5.8
Maximum	95	98	104	98
Mean	25	26	30	35
Standard deviation	14	14	16	17

Minimum alpha diversity increased logarithmically with increasing grain (Figure 4.5a). The relationship between the maximum alpha diversity and grain was nonlinear (hump-shaped) with diversity increasing between quarter and half degree grain, attaining a peak at one degree grain and then declining towards the coarser two degree grain (Figure 4.5b).

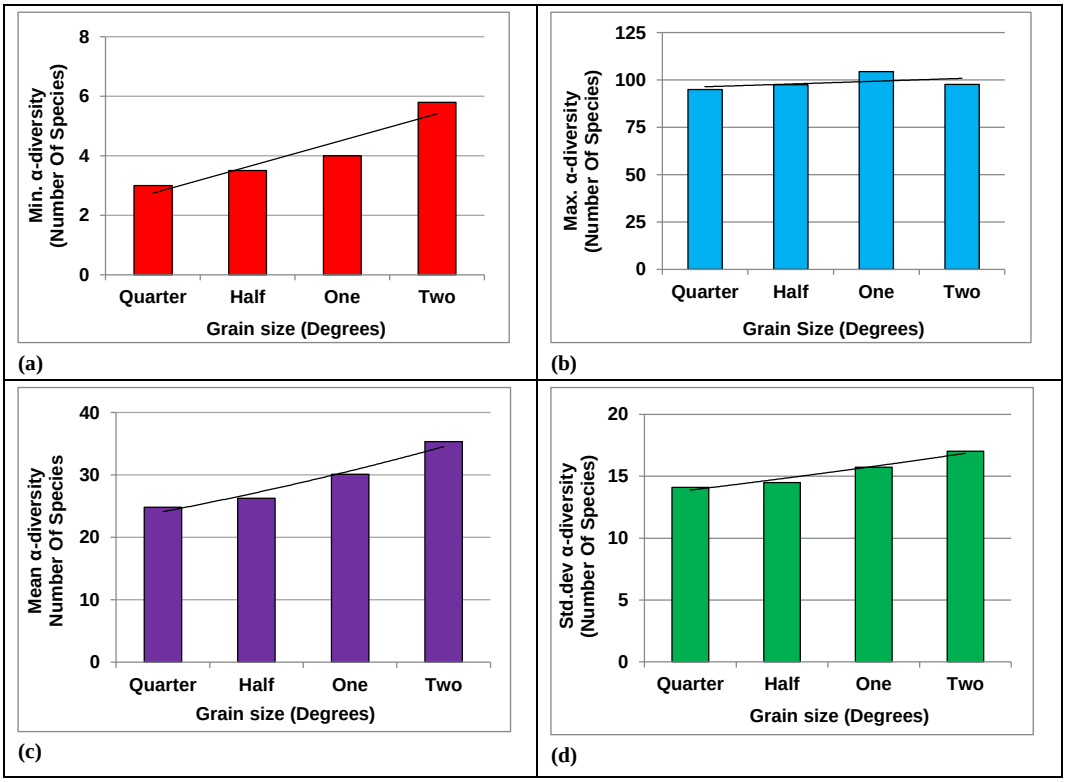


Figure 4.5 Variation in the (a) minimum (b) maximum (c) mean and (d) standard deviation of alpha diversity across grain.

Similar to the minimum parameter, mean alpha diversity (Figure 4.5c) increased as an power function of grain. The increase in mean alpha diversity was larger between the coarser grains (1º - 2º) than the finer ones (0.25º - 0.5º). The nonlinear increase in mean species richness is in keeping with the findings of several studies (Arita & Rodriguez 2002; Condit et al. 2002) and has been attributed to the nonlinear responses (i.e. ecological & historical) of biota to changes in environmental conditions (i.e. increased habitat heterogeneity) with increasing grain (Holt et al. 1999).

Lastly, the standard deviation of alpha diversity increased linearly as a function of spatial grain although the relationship remained constant (SD = 14 species) between the two finer grains. The increase in sample variance, though small, was unexpected because coarser spatial grains are typically associated with lower sample variance due to an increase in spatial smoothing here. These results are thus in direct conflict with the literature of Wu & Qi (2000) which suggests that aggregation decreases variance across grain due to spatial smoothing. However, Holt et al. (1999) note that variation in species richness is expected to be higher at coarser grains due to an increase in habitat heterogeneity.

4.4.2 **Beta diversity**

Table 4.2 and Figure 4.6 show the summary statistics extracted for beta diversity across spatial grain. The minimum beta diversity remained consistent across grain and never exceeded zero (Figure 4.6a). In contrast, maximum beta diversity increased linearly across grain (Figure 4.6b).

The increase in maximum beta diversity was greater between the larger grains (+0.19) than the finer grain (+0.06).

Table 4.2 Descriptive statistics of beta diversity across grain.

	Grain Size (degrees)			
Statistic	**Quarter**	**Half**	**One**	**Two**
Number of observations	23152	5943	1501	417
Minimum	0.00	0.00	0.00	0.00
Maximum	0.39	0.45	0.49	0.60
Mean	0.05	0.118	0.17	0.31
Standard deviation	0.050	0.06	0.08	0.10

Mean beta diversity (Figure 4.6c) increased as a stepwise linear function with increasing grain. These results are in agreement with those of Barton et al. (2013) who found biotic dissimilarity to increase with increasing grain, but contrast those of Keil et al. (2012) who found biotic dissimilarity to decrease with increasing spatial grain. These two studies however, measured changes in assemblage composition using two different indices of diversity and for different taxa (birds vs amphibians).

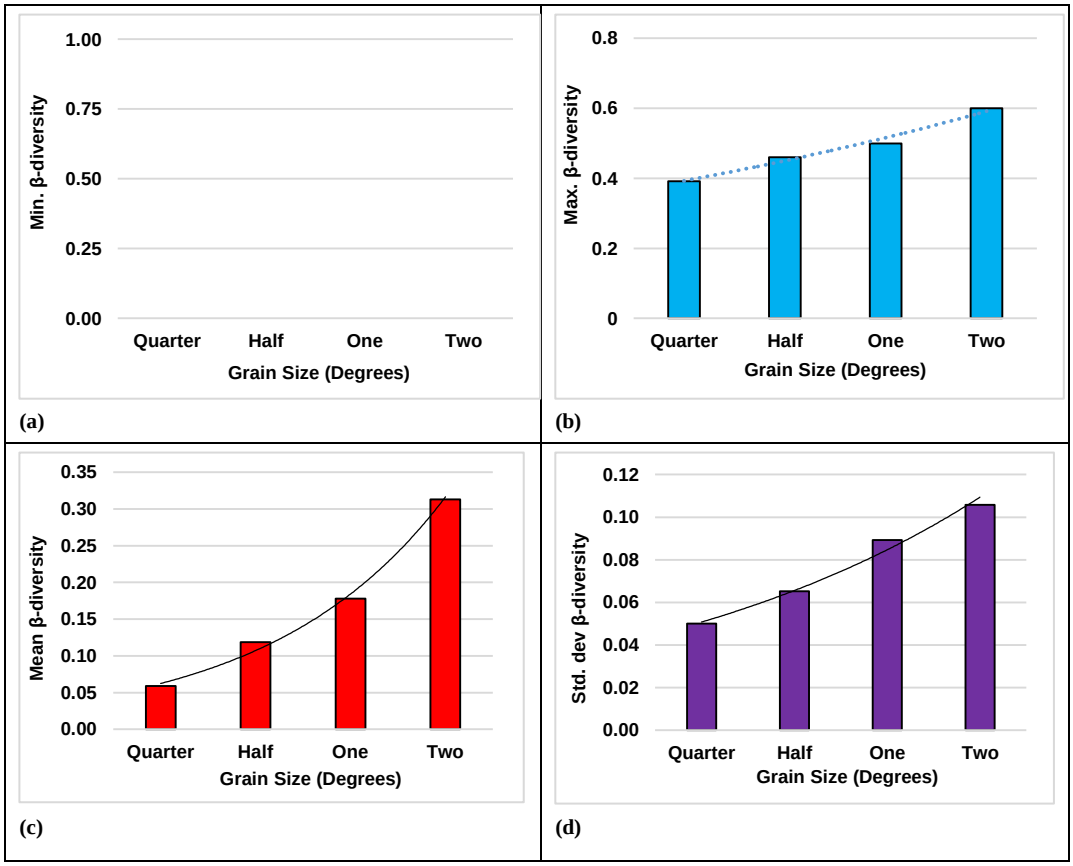


Figure 4.6: Variation in (a) minimum (b) maximum (b) mean and (d) standard deviation of beta diversity across grain.

This suggests that the scaling behaviour of beta diversity is intrinsically linked to both the taxa of interest and the index used to discriminate turnover. Surprisingly, the standard deviation (Figure

4.6d) of beta diversity also increased with increasing grain which suggests an increase in sample variance with increasing sampling grain. Fitting model trend lines to the data shows that the increase was nonlinear (exponential) in nature. This behaviour contrasts the findings of Wu & Qi (2000) who found that within sample variance of landscape metrics to decrease linearly as function of spatial grain. However, Holt et al. (1999) suggests that higher habitat heterogeneity at larger spatial grains increases the variance of beta diversity (i.e. larger grains stretch across more habitats).

Note: When regressed linearly as function of grain, only mean alpha diversity and beta diversity exhibited significant relationships, as determined by observing the p-value of the model slope, with grain.

4.5 GRAIN DEPENDENCY: PATTERN-PROCESSES RELATIONSHIPS

4.5.1 Alpha diversity

Table 4.3 shows the variation in regression parameters of the OLS and GWR models across spatial grain. Parameters associated with both the OLS and GWR model were grain variant. For both models, the model coefficient increased with increasing grain but remained significant ($p < 0.05$). Similarly, the significance and magnitude of the coefficients associated with the predictors remained consistent across grain for both models (Table 4.3). An exception however was TRange which presented a positive significant relationship with alpha diversity in the OLS model at half and one-degree grain but presented an insignificant contribution towards the model at two-degree grain.

Table 4.3: OLS and GWR model parameters across spatial grain. Signficance ($p < 0.05$) is indicated by an asterisk (*).

Model		**OLS**			**GWR**		
		Grain Size			**Grain Size**		
		Half	**One**	**Two**	**Half**	**One**	**Two**
	Intercept	-1.15*	-1.76*	-1.96*	1.21	3.34	3.51
Coefficient	**MAP**	0.16*	0.3*	0.60*	< 0.0001	< 0.001	< 0.001
	TRange	0.02*	0.08*	0.05	< -0.0001	< -0.0001	< -0.01
	NPP	0.15*	0.11*	0.13*	< 0.0001	< - 0.001	< - 0.001

Initially, a modelling error was suspected, however, subsequent data vetting to source potential erroneous analysis found the relationship to hold consistent. For this reason, the results were interpreted as is. Field et al. (2008) showed that variable significance associated with the species richness of birds fluctuated with spatial grain which they attributed to changes in the spatial autocorrelation lengths of the predictor and response variables. That is, a change in the

significance of the spatial dependence between predictor and response variable along a scale domain.

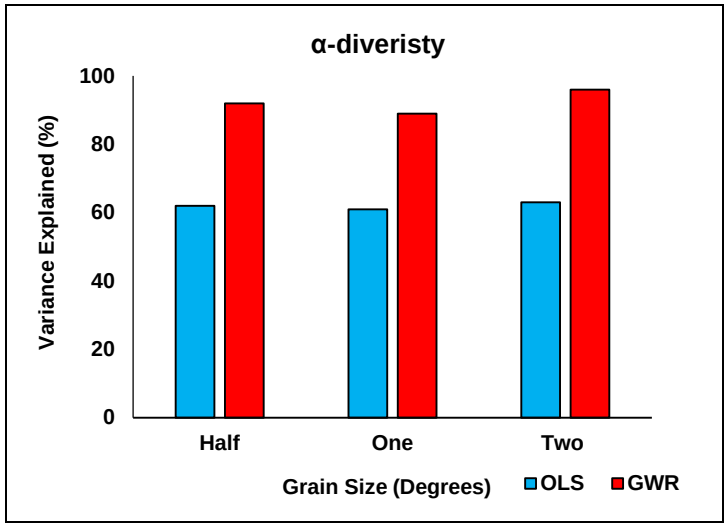


Figure 4.7: Variation explained in alpha diversity patterns by OLS and GWR models across spatial grain.

For both the OLS and GWR (Figure 4.7; Table 4.4) model fit, as explained by the adjusted R^2, remained relatively consistent. The best fit was recorded at the half-degree grain (R^2 = 0.62) followed by the two-degree (R^2 = 0.63). Model fit was slightly weaker at one-degree grain, with environmental correlates explaining only 61% of the variation in species richness (R^2 = 0.61).

Table 4.4: Selected model evaluation criteria across grain. *Asterisk indicates significant variables ($p < 0.05$).

	OLS			GWR		
Criteria	**Grain**			**Grain**		
	Half	**One**	**Two**	**Half**	**One**	**Two**
Adj. R^2	0.62	0.61	0.63	0.92	0.89	0.96
AIC	-3241.89	1587.94	332.80	-9629.44	-219.88	-148.57
Morans I	0.92	0.90*	0.85*	0.61*	0.81*	0.46*

For the OLS, in terms of model performance as indicated by the AIC values, the best performing model was recorded at the half-degree grain followed by the two-degree grain. The lower AIC scores at these two resolutions indicate improved model parsimony with respect to the number of variables in the model and total variance explained. The OLS model at one-degree resolution demonstrated the poorest model performance as indicated by the relatively high AIC score (AIC = 1587.94). For the GWR, AIC scores increased across grain indicating a decrease in model performance across grain. The grain dependency of model performance thus seems to be intrinsically linked to the statistical tool used to model the pattern-process relationships. The spatial dependence between model residuals, as measured by the Morans I value, decreased with increasing grain for the OLS model. This was expected, because as grain is increased, the spatial distance between sites is also increased, thereby leading to a disassociation in species co-

occurrence patterns due an increase in habitat complexity (i.e. encountering new species) and environmental dissimilarity. In contrast however, a hump-shaped relationship in the spatial dependence of model residuals was observed for the GWR model. This triangular behaviour was surprising because spatial dependence is considered to decrease linearly as a function of distance due to an increase spatial lag. However, studying the relationship between spatial autocorrelation and spatial grain, Qi & Wu (1996) found that although spatial autocorrelation generally decreased as a function of grain, random peaks and troughs still occurred across grain. They attributed this behaviour to scale thresholds: locations or regions along a scale domain where discrete changes or abrupt breaks in pattern-process interrelations occurs.

With the GWR, model fit also demonstrated a humpback behaviour across grain (Figure 4.7). The best model fit was obtained at two-degree grain ($R^2 = 0.95$) followed by half-degree grain ($R^2 = 0.91$). Similar to the OLS model, the one-degree grain presented the weakest fit ($R^2 = 0.89$) though the difference was relative small (6% < max R^2). The nature of model fit across grain demonstrated by the local GWR contrasts the findings of Foody (2004) which observed an linear reduction in model fit with coarsening grain. Foody (2004) however, used simple regression models to examine diversity-environment relationships (i.e. less complex) opposed to multiple regressions (more complex) used here which may explain the observed difference in model behaviour across grain.

Overall, the fit provided by the GWR models remained superior to that of the OLS models across spatial grain. In contrast to the OLS model, model performance for the GWR decreased across grain as indicated by the increasing AIC scores. The GWR also corrected best for spatial dependence, despite the peak in Morans I value at the one-degree grain, as shown by the decline in Moran's I value across grain.

As seen by Figure 4.8 and Figure 4.9, the spatial distribution of model residuals produced by the OLS and GWR models for alpha diversity remained consistent across grain.

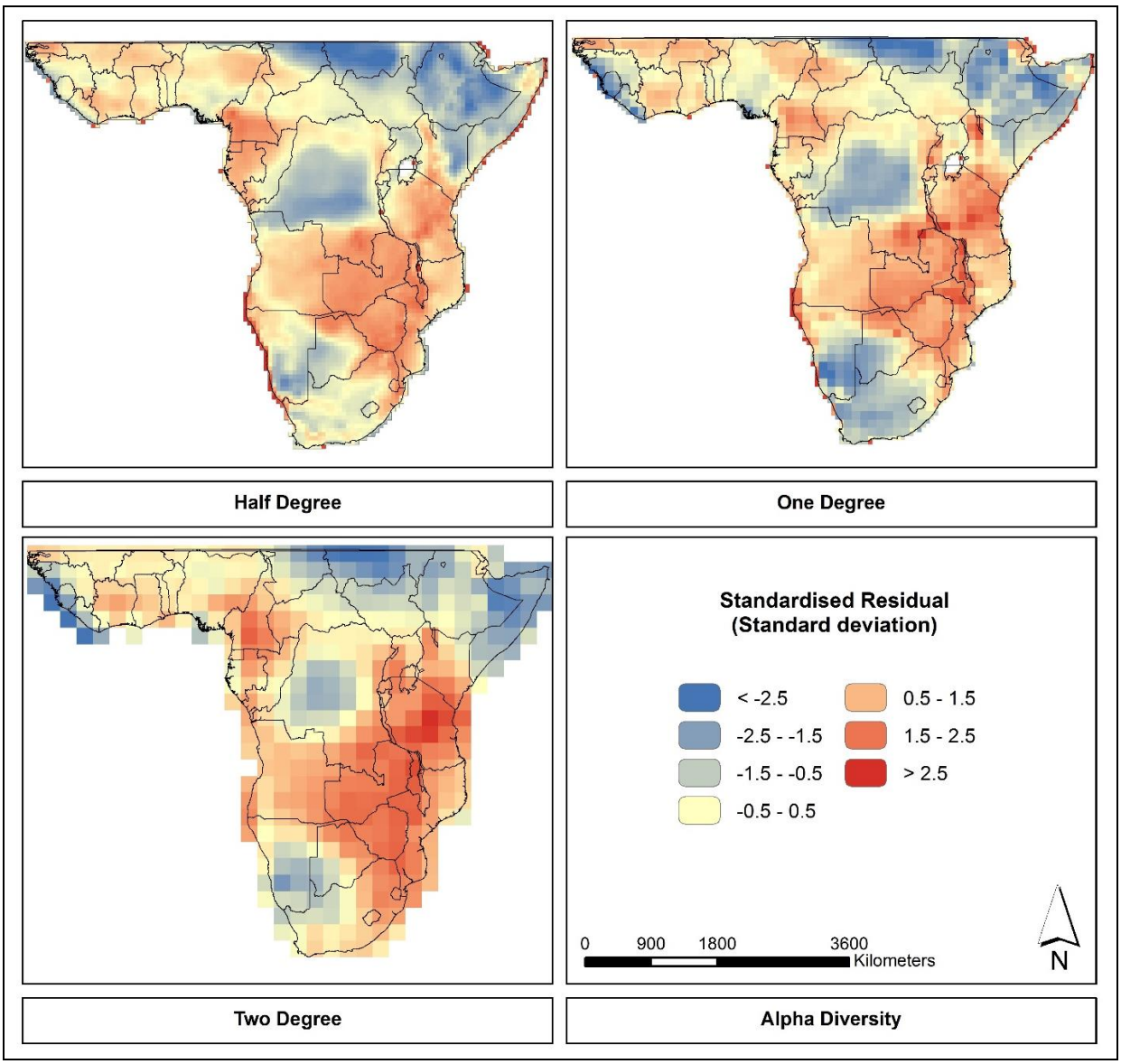


Figure 4.8: Spatial distribution of standardised residual of the GWR model for alpha diversity at (a) half-degree (n) one-degree and (c) two-degree resolution.

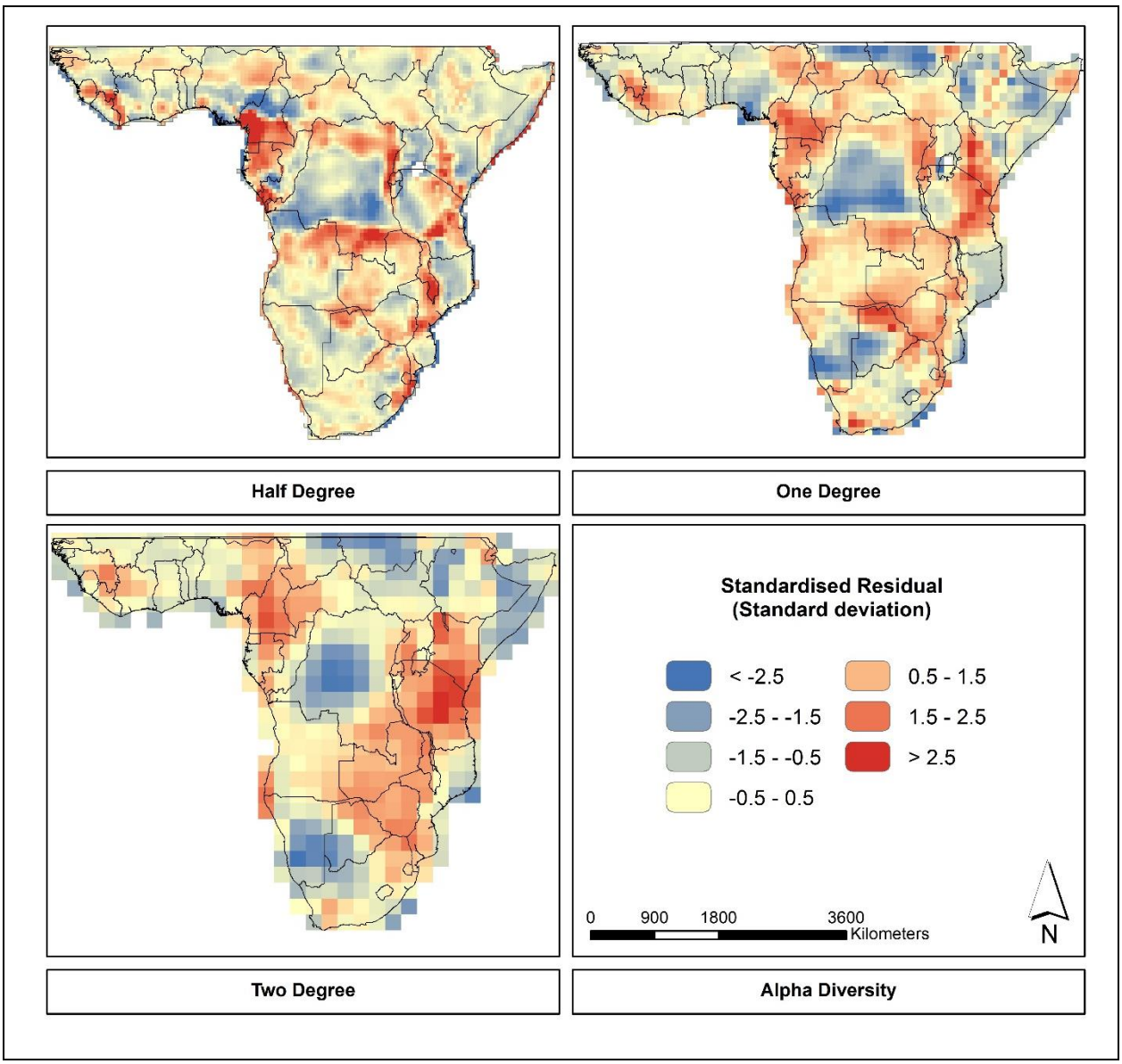


Figure 4.9: Spatial distribution of GWR standardised residuals for alpha diversity at (a) half-degree (b) one-degree and (c) two-degree spatial resolution.

This reaffirms the notion that at broad spatial grains the overarching spatial patterns persisting across the study extent remained consistent despite a loss in fine scale detail and spatial heterogeneity of the system.

4.5.2 **Beta diversity**

For the OLS regression, the significance and direction of model intercepts remained consistent across grain (Table 4.5) though the magnitude of these was grain variant. At all three grains, the model intercept was positive. The two-degree grain presented the steepest slope and the one-degree grain the gentlest. Similarly, the contribution of each predictor towards the model remained significant across grain. The contribution magnitude of each predictor, however, was grain-variant with MAT contributing the most towards the model at half-degree and one-degree grain but TRange contributing slightly more towards the model at two-degree grain. The nature and relationship between beta diversity and each predictor, however, remained consistent across

grain with all variables exhibiting a negative relationship with beta diversity except for topographic complexity (TOPSOD) which presented a positive relationship across grain.

Table 4.5: OLS and GWR model parameters for beta diversity across spatial grain. Significant variables are indicated by an asterisk* ($p < 0.05$).

Model		**OLS**		
		Grain Size		
		Half	**One**	**Two**
	Intercept	1.44*	1.35*	3.86*
Coefficient	**MAP**	-0.08*	-0.11*	-0.15*
	MAT	-0.41*	-0.25*	-0.52*
	TRange	-0.17*	-0.14*	-0.56*
	NPP	-0.03*	-0.01*	-0.11*
	TOPOSD	0.05*	0.04*	0.07*

The change in relationships between predictors and response variables across spatial grain observed here concur the findings of Levin (1992) who attributed this scale-variant behaviour to the variation in spatial autocorrelation length between patterns and pattern generating processes. Simply, this suggests that different processes are responsible for generating patterns at different scales. Furthermore, model performance for both the OLS and GWR was variable across grain (Table 4.6)

Table 4.6: Selected model evaluation criteria across grain. * Asterisk indicates significant variables ($p < 0.05$).

	OLS			**GWR**		
Criteria	**Grain**			**Grain**		
	Half	**One**	**Two**	**Half**	**One**	**Two**
Adj. R^2	0.12	0.15	0.20	0.54	0.48	0.52
AIC	1801.88	3309.27	287.49	78.12	2919.93	92.75
Moran I	0.67*	0.69*	0.66*	0.52*	0.61*	0.51*

For the OLS model, model fit as described by the adjusted R^2, increased with increasing grain (Table 4.6). In particular, at half-degree grain the model could explain only 12% of the variation in beta diversity to the 15% and 20% explained by the models at one-degree and two-degree grain respectively. The improved model fit demonstrated across grain may in part be due to increased spatial smoothing prevalent at coarser grains, which decreases the variance of the data consequently improving model fit. The improved model fit of data at coarser grains has been

noted by Goodchild & Quattrochi (1997) who attribute this functional behaviour to a decrease in sample variance at coarser grains.

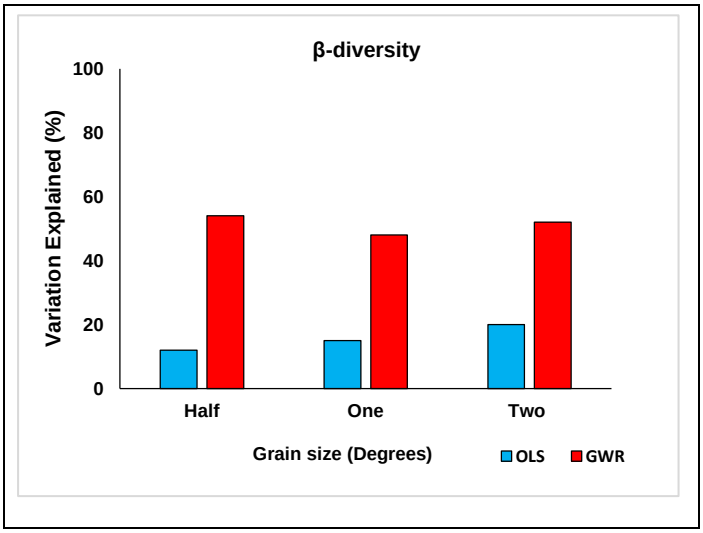


Figure 4.10: Variation explained in beta diversity patterns by OLS and GWR models across spatial grain.

Conversely, for the GWR, model fit decreased from 54% explained variance at 0.5º grain to 48% at 1º grain but then increased to 52% examined variation at 2º grain. This inconsistent behaviour in model fit corresponds to the findings of Rahbek (2005) who examined the relationship between avian diversity and elevation but found no distinct behavioural relationship across grain. The results, however, contrasts those findings of Field et al. (2008) who found the primacy of environmental variables in explaining diversity to increase with increasing grain. The decrease in model fit demonstrated by the GWR at one-degree grain may thus reflect an actual change or fluctuation in the beta diversity-environment relationship or be an artefact of the statistical analysis used to explore the relationship between the variables. Specifically, at coarser grains fewer sampling units are included in the statistical analysis for the GWR because of the increase in the size of the sampling unit (there are less neighbours to analyse).

Similarly, as indicated by the AIC and Moran's I scores, model performance and spatial dependence for both the OLS and GWR models increased between the half degree and one-degree grain and decreased between one-degree and two-degree grain (i.e. triangular behaviour). Again, the fluctuation in model performance and spatial dependence at one-degree grain may reflect an actual change or fluctuation in the local beta diversity-environmental correlate relationship or be an artefact of the statistical analysis used to explore the relationship between the variables.

Furthermore, checking the spatial distribution of the standardised residuals (Figure 4.11 & Figure 4.12) shows that though the underlying distribution of the data were affected by altering grain, the relative spatial distribution (where) of the model over and under estimates remained

consistent across grain. Specifically, both OLS and GWR models over-predicted in areas of high turnover and under-predicted in areas of low turnover across spatial grain.

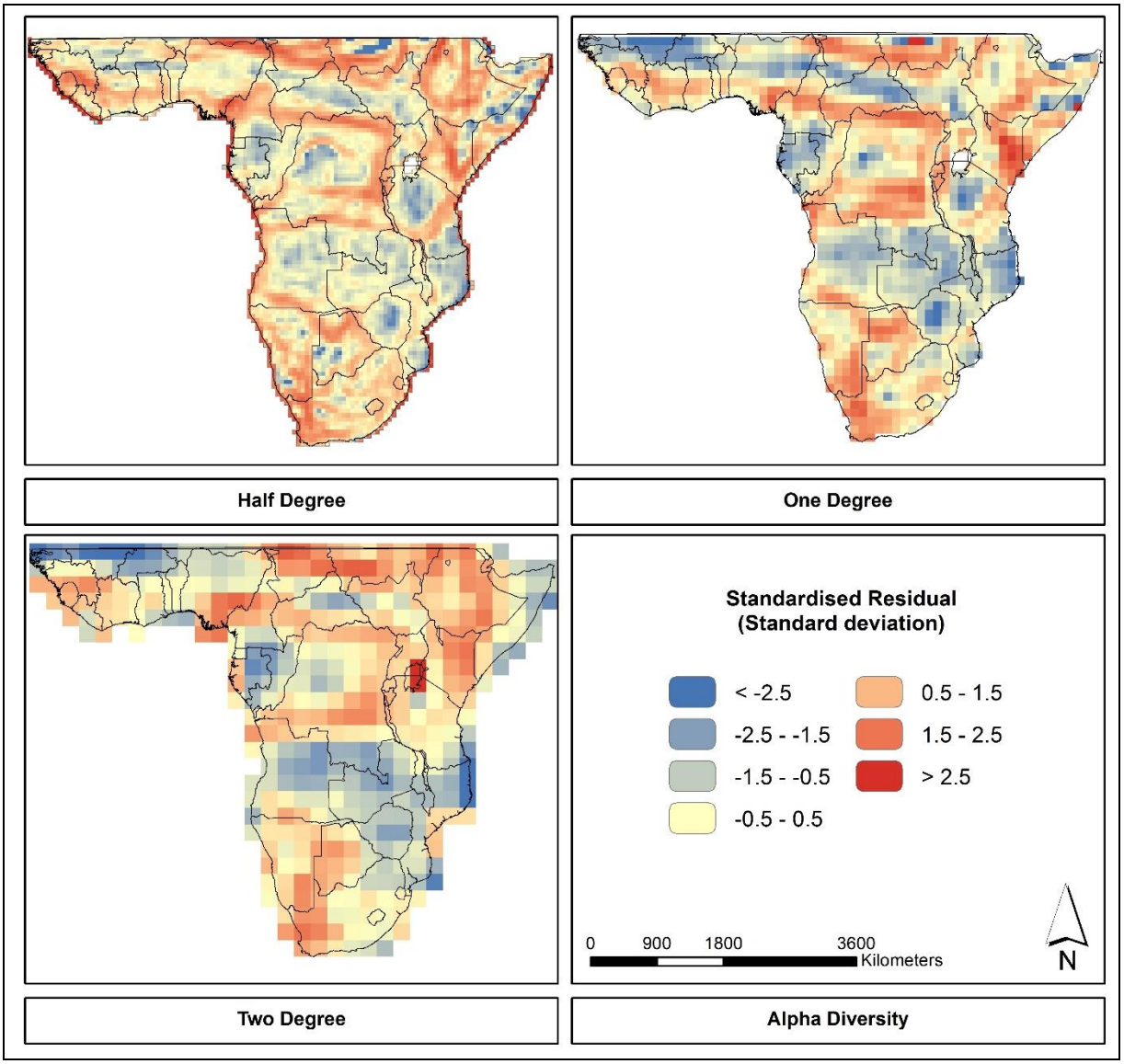


Figure 4.11: Spatial distribution of standardised residual of the OLS model for beta diversity at (a) half-degree (b) one-degree and (c) two-degree resolution.

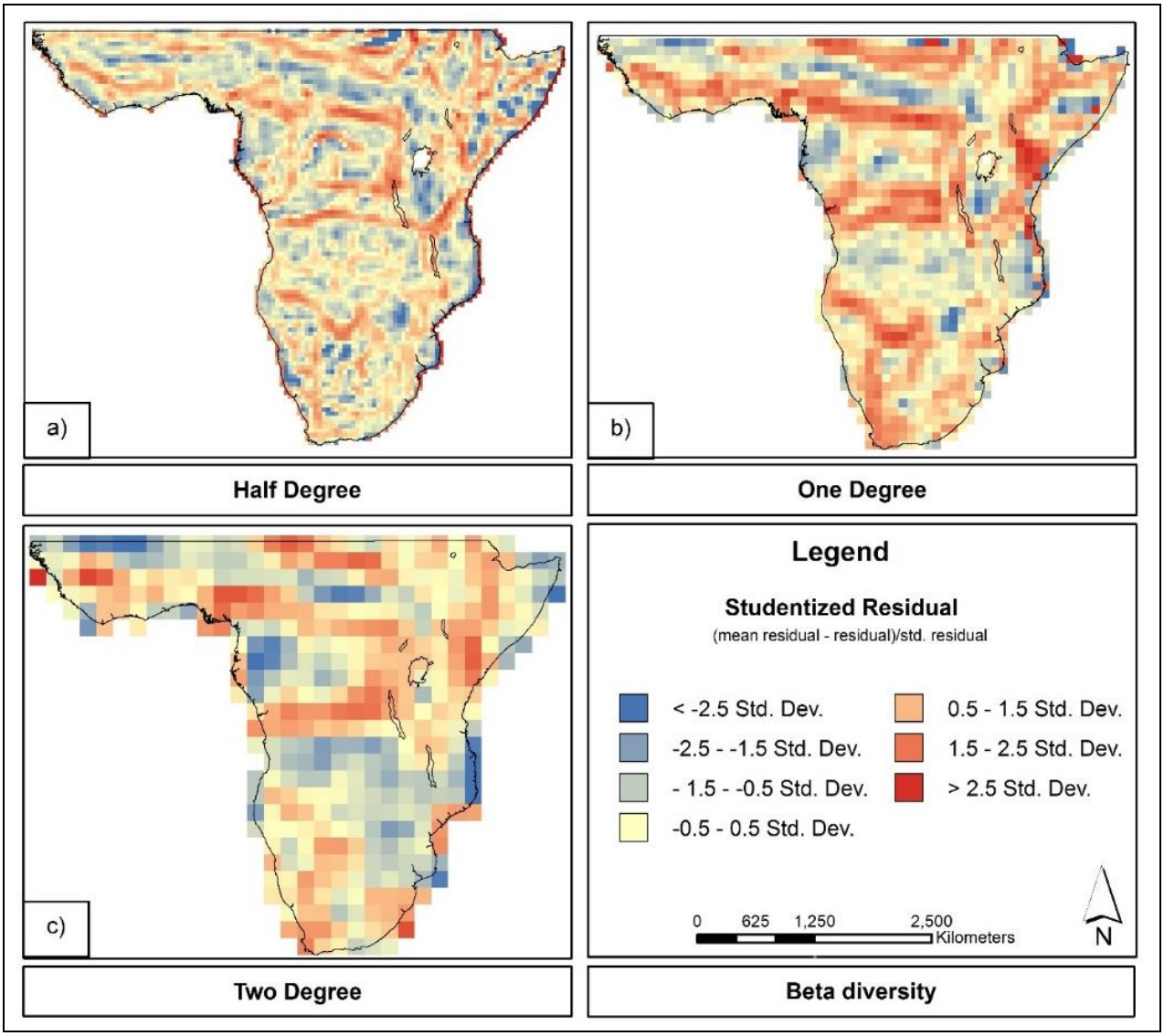


Figure 4.12: Spatial distribution of standardised residual of the GWR model for beta diversity at (a) half-degree (b) one-degree and (c) two-degree resolution.

4.6 DISCUSSION

Wu & Qi (2000) note that variables used to examine scaling effects generally exhibit three types of behaviour:

> **Type 1:** Predictable with linear, power law, or logarithmic function behaviour.
>
> **Type 2:** Predictable step-wise behaviour
>
> **Type 3:** Erratic responses exhibiting unpredictable and inconsistent scaling responses.

The same typification was used in this study to explain behaviour summary statistics of across grain.

4.6.1 Patterns and metrics of diversity are grain dependent

Results from the present study show that spatial patterns of diversity, both richness and turnover, are scale dependent. In particular, visual analysis show that patterns of diversity remain consistent across grain but that considerable fine scale detail is lost as grain site is progressively enlarged. In addition, greater degrees of spatial smoothing were associated with the coarser grains. Visually, this suggests a net loss in the spatial heterogeneity of a system with increasing grain. The smoothing observed at coarser grain reflects a percolation process (Hui 2009). That is, with an increase in grain, adjacent cells merge as a cluster thereby reducing the spatial heterogeneity of a given system. The observation of this percolation effect is important because it suggests a reduction in the spatial dependence between sampling units (i.e. species distributions) with an increase in grain.

The behaviour of alpha and beta diversity as a function of grain with constant extent were analysed in this study using summary statistics. All summary statistics showed strong scale dependence. Minimum alpha diversity increased with increasing grain while maximum alpha diversity demonstrated a hump-shaped behaviour across grain. The turning points on the humpback function suggest scale breaks: regions or location along a scale domain where the relationship between pattern and process changes (Hui 2009). Mean alpha diversity increased exponentially as a function of grain. The increase in species richness with increasing grain relates to the species-area hypothesis which posits that larger sampling areas allow for the detection of more species. The exponential increase in species richness across grain however contrasts the findings of Tokeshi (1993) who found species richness to increase as a power function across spatial grain. A power function is logical because regional species pools are expected to become saturated (i.e. reach equilibrium) as sampling becomes exhaustive. However,

both power and exponential functions exhibit linear behaviour if a log-log transformation is applied. A linear species-area curve in log-log space corresponds to the established literature (Adler & Lauenroth 2003). Furthermore, the standard deviation of alpha diversity increased with increasing grain. This result is in agreement with that of Hui (2009), which in a similar scale dependence study deduced a "rich get richer" effect in data when increasing grain size, but in contrast to that of Wu (2004) which found that aggregation decreases variance across grain. A decrease in variance was expected because spatial smoothing is expected to decrease within-sample heterogeneity at coarser grains. However, the results observed here indicate that the opposite might also be true particularly at large spatial extents.

With respect to beta diversity, the minimum beta diversity remained consistent across grain while the maximum beta diversity showed a linearly scaling relation i.e. Type 1 & 2 scaling relations. Surprisingly, mean beta diversity increased linearly as a function of grain. The increase in beta diversity with increasing grain are consistent with the physiology of the amphibian. Amphibians have low vagility and poor dispersal capacity. Thus, at coarser spatial grains the influence of geographical distance gains primacy in driving biotic dissimilarity between sites as amphibians cannot readily disperse to environmentally favourable sites. Keil et al. (2012) showed that geographic distance was the strongest determinant of beta diversity at coarser spatial grains. An alternate reason for higher beta diversity at coarser grains may be due to the progressive increase in environmental dissimilarity between sites at coarser grains. In particular, as the distance between two sites increases, so do the environmental conditions associated with those sites. As species are intrinsically adapted to their environments, greater geographical distances between sites result in steeper environmental gradients associated with those sites consequently resulting in greater changes in species found between those sites and hence, higher beta diversity. The results from regression models partially support this notion as the amount of variance explained by environmental variables were greater at coarser than finer scales. This, however, may also be attributed to an increase in spatial smoothing at coarser grains resulting in a decrease in data variance in data space (i.e. closer values in data space result in improved fit). The increase in beta diversity with increasing grain has been previously been documented by Navarro-sigüenza & Rodríguez (2014) for the amphibians of Mexico. In contrast to the patterns observed here, other studies (Arita & Rodriguez 2002; Mac et al. 2004; Wang et al. 2012) have reported higher beta diversity at finer spatial scales. Recently, Keil et al. (2012) reported a decrease in the biotic dissimilarity of European amphibians with increasing grain. The apparent contradiction of the results can be explained mainly by differences in conceptual terms and, in consequence, in the methods applied to measure beta diversity. That study used the Simpson index to the Sorensen index used here. Additionally, Arita & Rodriguez (2002) warn

that proportional indices (i.e. normalised measures), indices which are affected by richness gradients such as the Sorensen index, are influenced by the species-area relationship and thus will have higher beta diversity at coarser grains due to an increase in the number of species shared at coarser grain (the *a* parameter of the index increases). Furthermore, as with alpha diversity, the standard deviation of beta diversity increased with increasing grain which was unexpected as a decrease in within sample heterogeneity was expected to bring about a reduction in sample variance at coarser grains. The increase may in part be due to an increase in habitat heterogeneity at coarser grains. That is, larger sampling quadrats stretch over a greater variety of habitats (i.e. ecoregions) thereby including a greater variety of species and hence neighbouring quadrats possess greater beta diversity.

4.6.2 **Correlates of diversity**

4.6.2.1 Alpha diversity

Analysis of results in this study found alpha diversity-environments relationships, as determined through regression analysis, to be grain sensitive. This is in keeping with the findings of similar studies (Blackburn & Gaston 2002; Ewers & Didham 2008) and reaffirm the pervasive influence of scale on spatial analysis.

For alpha diversity, the significance and magnitude of environmental predictors varied as a function of grain. The magnitudes of NPP and TRange coefficients were consistent and erratic across grain, respectively. TRange was significant at the finest resolution of the study but not at the coarsest, supporting a changing pattern-process relationship across grain. Subsequently, the contribution of each predictor towards the model (i.e. OLS) also varied as a function of grain with the contribution of MAP, the largest contributor towards the model, increasing with coarsening grain. This may indicate that the primacy of precipitation in determining amphibian diversity is greater at larger than finer spatial scales or simply be an artefact of spatial smoothing reducing within sample variance at large spatial grains.

Model evaluation criteria were also grain-variant. Model performance, as indicated by the AIC scores, decreased with increasing grain size. Surprisingly, model fit remained relatively consistent across grain (± 62%). For the GWR, model fit increased with increasing grain, with the exception of the one degree grain where a decline was observed. Regardless of this, the fit was consistently high, irrespective of the grain of analysis; this and suggests that amphibian diversity is highly coupled to present day environmental conditions at various spatial scales. The consistency in model fit across grain corresponds to those findings of Buckley & Jetz (2007) which found the relationship between amphibian diversity and environmental correlates to remain consistent across grain. However, these findings contrast those findings of several other

studies (Cornell & Lawton 1992; Lennon et al. 2001; Ewers & Didham 2008) which found the variation of alpha diversity explained by environmental predictors for other taxa to increase with increasing grain. Blackburn & Gaston (2002) note that the discrepancies in variation explained by environmental correlates across grain between studies may, in part, reflect actual changes pattern-process interrelationships or be attributed to conceptual differences in the methodological framework followed between studies. The discrepancies in results further show the perplexing influence of spatial scale on the detection and quantification of pattern-process relationships.

Furthermore, spatial dependence as measured by the Morans I value, also declined with increasing grain. The decline in spatial dependence across grain reflects the notorious distance-decay of similarity phenomena which posits that biotic similarity between sites decreases as a function of the distances between those sites owing to progressively dissimilar environmental conditions associated with those sites.

4.6.2.2 Beta diversity

Few studies have examined the grain dependency of beta diversity and its determinants and thus there is no formal theory predicting exactly how grain size should affect the environmental correlates of beta diversity. Keil et al. (2012) investigated the relationship between amphibian beta diversity, environmental distance and geographic distance across grain and showed that geographic distance consistently explained the most variation in amphibian diversity across grain. They further showed that environmental correlates explained more variation in beta diversity at coarser than finer grains. Similarly, Calderón-Patrón et al. (2013) showed that geographic distance explained more variation in beta diversity at coarser than finer grains.

Results from the study concur the grain dependency of determinants of beta diversity. Using a global OLS and local GWR to regress environmental correlates with beta diversity, the study showed that environmental determinants of beta diversity are grain dependent. Both model parameters and diagnostics were found to covary with spatial grain. As seen by the increase in the adjusted R^2 value, the OLS model suggests that environmental factors are stronger predictors of beta diversity at coarser than finer grain which corresponds to the findings of Keil et al. (2012). The increase in variation explained by environmental correlates at coarser grains has been accredited to spatial smoothing (Mac Nally et al. 2004). Spatial smoothing or averaging indicates a reduction in within sample variance whereby variance of data the data is reduced, thereby increasing the overall model fit. Spatial smoothing is thus an artefact of the spatial analysis. In contrast, the variance explained by the local GWR model remained relatively consistent across grain which may indicate that the observed relationship between beta diversity and environmental correlates across grain also depends on the statistical method used to model

the data. Nonetheless, the consistency in variance explained by the GWR across grain contrast the findings of Foody (2004) which showed that the GWR decreased with increasing grain.

Overall, spatial dependence in the model residuals declined with increasing grain which may indicate a disaggregation in the spatial association of species co-occurrence patterns.

4.6.3 **Implications for conservation biology**

Effective conservation strategies require an understanding of how processes operating across multiple scales influenc anuran distributions. Many conservation and ecosystem management strategies are informed by studies conducted at a single, fine grain of analysis (i.e. analytical scale), explicitly assuming inferences made at this scale holds at another. In addition, they assume that because ecosystems follow a hierarchal structure, scaling responses are likely to be linear.

Hui (2011) notes that such inferences are erroneous because the analytical scale and measurement scale reflect only a single portion of the characteristic scale of a phenomenon. That is, because species distributions are regulated by a variety of intertwined abiotic and biotic processes operating at characteristic different scales, information being picked up is diluted by the measurement scale. Furthermore, the interaction between the processes themselves and the response of taxa to these processes may not be linear.

The results from the present study, like those of several others (Lennon et al. 2001; Mac Nally et al. 2004; Nthiwa 2011), assert support to the aforementioned notion and show that the scaling response of anurans diversity and the relationship it exhibits with environmental processes, are not necessarily linear in nature, but exhibit a variety of complex scaling responses. The nature of the responses is influenced by the; physiology of anurans, the response of anurans to interaction between abiotic and biotic processes, the statistical methods used to make inferences as well as the scale of analysis grain of analysis. Thus, although species richness (i.e. preserve where there are more) and turnover (preserve where there is change, rare) are both useful to inform conservation policies, many current conservation strategies incorporating either diversity component remain misguided in their purpose to conserve because of the scale problem.

If adequate and robust conservation strategies are to be drawn up, spatial scale, effectively a component of the MAUP, needs to be considered.

4.7 CONCLUSION

The aim of this chapter was to evaluate the relationship between amphibian diversity patterns and spatial scale. The overarching research objectives were to observe how patterns of

amphibian diversity, metrics describing amphibian diversity and environmental correlates of diversity vary as a function of coarsening spatial grain. Visual results showed that patterns of diversity are grain sensitive with the overarching patterns of diversity remaining consistent but at the expense of fine scale detail. Metrics describing diversity were also grain-variant, varying as linear, exponential hump-shape and power functions across spatial grain. Using both spatial (GWR) and aspatial (OLS) regression techniques, diversity-environment relationships were also found to be sensitive to spatial scale. The nature, strength and model coefficients also varied as function of spatial grain.

Given the scale effects, this study concludes that spatial scale comprises a critical, concept in ecology and understanding the multiplicity relationship it exhibits with landscapes and ecosystems is imperative for understanding the mechanisms behind ecological-pattern-processes. Therefore, if we are truly to unlock and understand biological patterns, the issue of spatial scale and the scaling effect it imposes on landscapes and ecosystems must be addressed. The urgency to understand this imposition becomes ever more paramount as amphibian communities and other biota suffer unprecedented declines.

CHAPTER 5: SUMMARY AND SYNTHESIS

In this chapter, the evaluation and conclusions of the research findings are made. The conclusions made here follow the structure of the thesis objectives. Consequently, the first section revisits and provides an evaluation of each research objective. Thereafter, the conclusions of the study are drawn, while the last section offers recommendations and suggestions for further research.

5.1 REVISITING OBJECTIVES IN CONTEXT OF THE RESULTS

The overarching aim of this study was to examine the relationship between anuran diversity (richness and turnover) and the environment and to observe how spatial scale influences this relationship. This aim was to be realised through the achievement of six research objectives. In the succeeding sections, each objective is revisited and then related to the most pertinent findings of the study to evaluate the attainment of each research objective.

5.1.1 **Map the spatial distribution of alpha and beta diversity across the Afrotropics.**

The first research objective was to map the spatial distribution of anuran alpha and beta diversity across the study extent. Both diversity components were mapped using the newly developed zeta diversity partitioning function. The first order zeta component (ζ_1) was used to derive alpha diversity, while the Sorensen index, derived by manipulating first order and second order zeta components, was used to compute species turnover (β-diversity). The spatial distribution of alpha and beta diversity produced in the present study using zeta diversity correspond to established literature (Buckley & Jetz 2007, 2008). Zeta diversity partitioning, as an incidence metric, thus provides a feasible means to map broad scale biodiversity patterns and its use in capturing multi-assemblage structures should be further explored.

5.1.2 **Examine the relationship between diversity and the environment**

Multiple regression analysis was used to determine the relationship between diversity and the environment. In the regressions, alpha and beta diversity were the response variables and the environmental correlates were the predictors. The regression analysis comprised one aspatial global model, the Ordinary Least Squares (OLS) regression and one local spatial model, the Geographically Weighted Regression (GWR). The global OLS model was used to examine the general relationship between diversity and the environment across the study extent while the local GWR model was used to examine the spatial variability in the relationships (i.e.

nonstationarity) and model parameters across the study extent as well as to account for spatially dependent relationships. The OLS and GWR models were evaluated and compared using the adjusted R^2, AIC, and Morans I (residual) values associated with each model. The adjusted R^2 was used to explain model fit, and the AIC score was used to determine how well the model performed with respect to the number of variables included in the model to the variance explained by the model. The Morans I test was used to grade the spatial autocorrelation in the residuals of each model.

The OLS model showed that environmental variables could explain 62 % of the variation in alpha diversity which was increased to 92% using the local GWR model. The substantial variation explained by the regression models suggests and reaffirms that contemporary environmental conditions have played a major role in the structuring amphibian richness. Overall, mean annual precipitation was the single strongest predictor explaining 51% of the variation in the amphibian richness across the Afrotropics. The primacy of precipitation in determining richness is consistent with the physiology of anurans who require water throughout their life cycle (i.e. laying eggs, survival). The unexplained variation in the regression model, however, suggests that historical and ecological factors have also been influential in the spatial organisation of contemporary species distributions and diversity patterns.

For beta diversity, the OLS model showed that environmental correlates could explain only 12% of the variation in biotic dissimilarity across extent which was increased to 54% using the local GWR. Intuitively, this implies that beta diversity is locally associative and influenced by local factors as opposed to more globally associative ones. The substantive unexplained variance from both regression models indicate missing variables unaccounted for in the regression models, model misspecification, or not considering historical (i.e. evolution, speciation) or ecological (i.e. competition, interaction) factors.

Furthermore, plotting GWR parameters such as the coefficient of each predictor and local r^2 values, which changed in magnitude, showed that the spatial relationship between diversity and the environment were variable across the study extent. The spatial patterns in the parameter estimates reveal the danger of using a single estimate for a parameter derived from a global model for local variable relationships. The single estimate for each parameter derived from the global regression technique fails to represent the relationship of beta diversity (i.e. turnover) and current environmental variables used here across most of the region and the OLS, in relation to the GWR, has poor descriptive and predictive power. The

GWR can thus suffice as a useful alternative or supplementary technique to model diversity and should be investigated.

5.1.3 Grain dependency of diversity: Patterns

The third research objective was to determine whether diversity patterns were scale dependent and to what extent. This was achieved by hierarchically nesting three sets of sampling quadrats covering the study extent, and then determining alpha and beta diversity by means of zeta diversity partitioning. Visual analysis showed that patterns of alpha and beta diversity were grain-variant. General patterns of diversity remained consistent across grain but fine scale detail was lost in favour of smoothing across grain.

5.1.4 Grain dependency: Underlying distribution

For objective four the grain dependency of diversity metrics was investigated using summary statistics. At each scaling interval the minimum, maximum, mean and standard deviation of diversity were recorded to compare changes in diversity across grain. All summary statistics were found to be sensitive to spatial grain. Both mean alpha and beta diversity increased nonlinearly as a functions of spatial grain. The minimum and maximum variables displayed consistent, and nonlinear hump-back scaling behaviour while variance increased linearly across grain. The inconsistency reinforces the notion that spatial processes exhibit different pattern-processes relationships at different spatial grains.

5.1.5 Grain dependency: Determinants of diversity

The final research objective was to examine the relationship between the environmental determinants of diversity across spatial grain. This was achieved by regressing diversity against environmental correlates and then recording model parameters, such as model fit and coefficient slopes, at each grain interval. Diversity-environment relationships were found to be grain dependent. Model parameters behaved variably across grain and were model specific. For example, the variance explained by environmental correlates in the OLS model but was more variable in local GWR model. For beta diversity, the variance explained by the environmental correlates increased linearly with increasing grain for the OLS model while the behaviour of the GWR was triangular (hump-back) across grain. For both alpha and beta diversity, coefficient slopes of predictors and model intercepts produced by the regression models were also grain variant. Overall, most scaling metrics displayed type 1 and type 2 scaling behaviour which confirmed that emergent spatial patterns are the outcome of different

processes operating at multiple spatial scales and thus a single scale of analysis can lead to misleading or even erroneous conclusions.

5.2 CONCLUSION

In conclusion, several deductions can be made from this study. First, zeta diversity partitioning provides a new and exciting means to measure and map incidence based broad scale diversity patterns across a range of spatial scales. Second, regional climate and physiography exhibit significant relationships with diversity but are more highly correlated with species richness than turnover. Third, historical and ecological factors, such as dispersal, have most probably also been influential in structuring species assemblages. Fourth, species distributions are spatially dependent and their interaction with the current environment is non-stationary across space. These relationships thus need to be investigated using spatially explicit techniques, such as the GWR, which accommodate for these phenomena. Fifth, spatial patterns, spatial data, and pattern-process relationships are all scale sensitive, exhibiting complex and nonlinear scaling relations across spatial grain which reflects their complex spatial organisation. It is thus important to note the scale of analysis as results obtained and inferences made are scale specific. Lastly, the impact of the Modifiable Aerial Unit Problem must always be addressed and considered in ecological and spatial analysis. Thus if were are truly to unlock patterns and pattern generating processes the issue of spatial scale must be addressed.

If we are to truly understand how biological patterns are generated and how broad scale processes interact to generate these patterns, we need to investigate and gain understanding of the interplay between pattern and process across multiple spatial scales.

5.3 LIMITATIONS AND RECOMMENDATIONS

The study suffers from several limitations pertaining primarily to the collected data and data analysis. First, presence-absence data reflect idealised distribution of species and thus spatial patterns produced from such data are a property of the sampling effort and may not reflect realised distribution of species. Though this limitation is beyond the control of the researcher, it is nonetheless important to note.

Second, the results show that zeta diversity partitioning is an effective way to measure and discriminate broad scale biodiversity patterns. Its use for future broad-scale ecological analysis is thus highly recommended.

Third, the Sorensen index comprises only one measure of beta diversity as turnover. The results produced here are thus interpretable for this index only. Due to the plethora of indices used to measure beta diversity and the lack of the overall consensus as to what exactly

species turnover is, the results rendered by the Sorensen index, as in the case of this study, are thus index-specific. For this reason, making use of multiple indices used to measure, scale and relate beta diversity to environmental processes are recommended and may provide greater insight into the scaling nature of beta diversity.

Fourth, our results show that diversity-environmental relationships are non-stationary across geographic space. Parameters derived from global models are thus useful to deduce the overall or mean relationship between diversity and current environmental condition across the entire study extent but local models, such as the GWR, provide more locally abundant and spatially aware information such as areas where there are deviations in mean parameters.

Fifth, in terms of the scaling component of the study, the use of only three grains sizes may be insufficient to infer pattern-process scaling relations. Initially, the researcher had envisioned making use of at least five sampling grains, ranging in size from a quarter-degree to four-degree resolution to observe the effect of altering spatial grain on the patterns and drivers of diversity. However, several problems, mainly statistical and computational in nature, were encountered using those two sets. For example, the quarter degree grid set consistently suffered from memory and modelling errors, owing to the large number of sampling quadrats ($n = \pm 25000$) while the four-degree grid set comprised too few sampling units ($n < 30$) to suffice for sufficient regression analysis. For these reasons, the two sets were regrettably omitted from empirical analysis. It is thus recommended that future research aims to test hypotheses across a greater number of scales (grains).

Lastly, the study was severely hindered by a lack of computational power. For example, running GWR models, which are notorious for being computationally intensive (Dormann et al. 2007; Kissling & Carl 2008; Fischer & Getis 2010), proved challenging particularly for the fine scale grids (i.e. 0.5^{o}). GWR are thus recommended only for smaller spatial extents, large spatial grains or in situations where sufficient and adequate computational power is available to the researcher. Again, this limitation falls beyond the control of the researcher but is nonetheless important to note.

REFERENCES

Adler P & Lauenroth W 2003. The power of time: Spatiotemporal scaling of species diversity. *Ecology Letters* 6: 749–756.

Alford R & Richards S 1999. Global amphibian declines: A problem in applied ecology. *Annual Review of Ecology and Systematics* 30: 133–165.

Allen AP, Brown JH & Gillooly JF 2002. Global biodiversity, biochemical kinetics, and the energetic-equivalence rule. *Science* 297: 1545–1548.

Almeida-Neto M, Frensel DMB & Ulrich W 2012. Rethinking the relationship between nestedness and beta diversity: A comment on Baselga (2010). *Global Ecology and Biogeography* 21: 772–777.

Anderson MJ, Crist TO, Freestone AL, Sanders NJ 2011. Navigating the multiple meanings of β-diversity: A roadmap for the practicing ecologist. *Ecology Letters* 14:19–28.

Anselin L 1989. What is special about spatial data? Alternative perspectives on spatial data analysis. *UC Santa Barbara: National Center for Geographic Information and Analysis* 57: 63–77.

Anselin L 2002. Under the hood Issues in the specification and interpretation of spatial regression models. *Elsevier* 27: 247–267.

Araújo MB 1996. Mapping the Boundaries: variation in the transition zones to the European Mediterranean Region.

Arita HT & Rodriguez P 2002. Geographic range, turnover rate and the scaling of species diversity. *Ecography* 25: 541–550.

Aronson RB, Macintyre IG, Precht WF, Murdoch TJT, Wapnick CM, Monographs SE, May N 2002. The expanding scale of species turnover events on coral reefs in Belize. *Ecological Monographs* 72: 233–249.

Astorga A, Death R, Death F, Paavola R, Chakraborty M, Muotka T 2014. Habitat heterogeneity drives the geographical distribution of beta diversity: The case of New Zealand stream invertebrates. *Ecology and Evolution* 4: 2693–2702.

Barton PS, Cunningham SA, Manning AD, Gibb H, Lindenmayer D, Didham R 2013. The spatial scaling of beta diversity. *Ecological Sounding* 22: 639–647.

Barwell LJ, Isaac NJB & Kunin WE 2015. Measuring β-diversity with species abundance data. *Journal of Animal Ecology* 84: 1112–1122.

Baselga A 2010. Partitioning the turnover and nestedness components of beta diversity. *Global Ecology and Biogeography* 19:134–143.

Baselga A 2013. Multiple site dissimilarity quantifies compositional heterogeneity among several sites, while average pairwise dissimilarity may be misleading. *Ecography* 36:124–128.

Baselga A, Gomez-Rodrıguez C & Lobo JM. 2012. Historical Legacies in World Amphibian Diversity Revealed by the Turnover and Nestedness Components of Beta Diversity. *pLOS* 7: e32341 – e32341.

Beale CM, Jack J, Yearsley JM, Brewer J & Elston DA 2010. Regression analysis of spatial data. *Ecology Letters* 13: 246–264.

Beguería S & Pueyo Y 2009. A comparison of simultaneous autoregressive and generalized least squares models for dealing with spatial autocorrelation. *Global Ecology and Biogeography* 18: 273–279.

Bini L, Alexandre J, Diniz-Filho F, Thiago F, Ragel M & Thomas S 2009. Coefficient shifts in geographical ecology: An empirical evaluation of spatial and non-spatial regression. *Ecography* 32:193–204.

Bishop TR, Robertson MP, Rensburg BJ Van & Parr CL 2015. Contrasting species and functional beta diversity in montane ant assemblages. *Journal of Biogeography* 42:1776–1786.

Blackburn TM & Gaston KJ 2002. Scale in macroecology. *Global Ecology & Biogeography* 11: 185–189.

Blaustein AR & Johnson PT 2003. The complexity of deformed amphibians. *Frontiers in Ecology and the Environment* 1: 87–94.

Blaustein AR, Romansic JM & Kiesecker JM 2003. Ultraviolet radiation, toxic chemicals and amphibian population declines. *Diversity and Distributions* 9: 123– 140.

Brien RM 2007. A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity* 41: 673–690.

Brown JH & Kodric-Brown A 1977. Turnover rates in insular biogeography: Effect of immigration on extinction. *Ecology* 58: 445–449.

Buchanan GM, Nelson A, Mayaux P, Hartley A & Donald PF. 2008. Delivering a global, terrestrial, biodiversity observation system through remote sensing. *Conservation Biology* 23: 499–502.

Buckley LB & Jetz W 2007. Environmental and historical constraints on global patterns of amphibian richness. *Proceedings. Biological sciences/The Royal Society* 274: 1167–1173.

Buckley LB & Jetz W 2008. Linking global turnover of species and environments. *Proceedings of the National Academy of Sciences of the United States of America* 10: 17836-17841.

Burgess ND, Rahbek C, Larsen FW, Williams P & Balmford A 2002. How much of the vertebrate diversity of Sub-Saharan Africa is catered for by recent conservation proposals? *Biological Conservation* 107: 327–339.

Burkle LA, Myers JA & Travis Belote R 2016. The beta-diversity of species interactions: Untangling the drivers of geographic variation in plant-pollinator diversity and function across scales. *American Journal of Botany* 103: 118–128.

Cade BS & Noon BR 2003. A gentle introduction to quantile regression for ecologists. *Frontiers in Ecology and the Environment* 1: 412–420.

Calderón-patrón JM, Moreno CE, Pineda-lópez R & Sánchez-Rojas G 2013. Vertebrate dissimilarity due to turnover and richness differences in a highly beta-diverse region: The role of spatial grain size, dispersal ability and distance. *PLoS ONE* 8: 1–10.

Carey C, Heyer WR, Wilkinson J, Alford RA, Arntzen JW, Halliday T, Hungerford L, Lips KR, Middleton EM, Orchard SA & Rand AS 2001 Amphibian declines and environmental change: Use of remote-sensing data to identify environmental correlates. Conservation Biology 15: 903–913.

Cayuela L, María J, Benayas R, Justel A & Salas-Rey J 2006. Modelling tree diversity in a highly fragmented tropical montane landscape. *Global Ecology and Biogeography* 15: 602–613.

Chapin JP 1923. Ecological aspects of bird distribution in tropical Africa. *The American Naturalist* 57: 106–125.

Charlton M & Fotheringham S 2009. Geographically Weighted Regression. *The Statistician* 47: 431–443.

Chave 2013. The problem of pattern and scale in ecology: What have we learned in 20 years? *Ecology Letters* 1: 4–16.

Chejanovski ZA & Wiens JJ 2014. Climatic niche breadth and species richness in temperate tree frogs. *Journal of Biogeography* 41: 1936–1946.

Chen S, Jiang G, Zhang J, Li Y & Qian H 2011. Species turnover of amphibians and reptiles in eastern China: Disentangling the relative effects of geographic distance and environmental difference. *Ecological Research* 26: 949–956.

Clarke A & Gaston KJ 2006. Climate, energy and diversity. *Proceedings: Biological sciences/The Royal Society* 273: 2257–2266.

Colwell RK 1988 Biodiversity: Concepts, patterns, and measurement, communities and ecosystems. *The Princeton guide to ecology* 12: 257–264.

Condit R, Pitman N, Egbert G, Leigh J, Chave J, Terborgh J, Foster R, Nunez V, Aguilar V, Valencia R, Villa G, Muller-Landau H, Losos E & Hubbel H 2002. Beta-diversity in tropical forest trees. *Science* 295: 666–669.

Cornell HV & Lawton JH 1992. Species interactions, local and regional processes, and limits to the richness of ecological communities: A theoretical perspective. *Journal of Animal Ecology* 61: 1–12.

Currie DJ 1991. Energy and large-scale patterns of animal- and plant-species richness. *The American Society of Naturalists* 137: 27–49.

Cushman S, McGarigal K & Neel M 2008. Parsimony in landscape metrics: Strength, universality, and consistency. *Ecological Indicators* 8:691–703.

Czarnota J, Wheeler DC & Gennings C 2015. Evaluating Geographically Weighted Regression models for environmental chemical risk analysis. 14: u117–127.

Davidar P, Puyravaud JP & Leigh EG 2005. Changes in rain forest tree diversity, dominance and rarity across a seasonality gradient in the Western Ghats, India. *Journal of Biogeography* 32:493–501.

De Klerk HM, Crowe TM, Fjeldså J & Burgess ND 2002. Biogeographical patterns of endemic terrestrial Afrotropical birds. *Diversity and Distributions* 8: 147–162.

Di Virgilio G, Laffan SW & Ebach MC 2012. Fine-scale quantification of floral and faunal breaks and their geographic correlates, with an example from south-eastern Australia. *Journal of Biogeography* 39: 1862–1876.

Di Virgilio G, Laffan SW, Ebach MC & Chapple DG 2014. Spatial variation in the climatic predictors of species compositional turnover and endemism. *Ecology and Evolution* 4: 3264–3278.

Diniz-filho AF, Barberi M & Lima-ribeiro MDS 2010. Climate stability and the current patterns of terrestrial vertebrate species richness on the Brazilian Cerrado. *Quaternary International* 222: 230–236.

Diniz-filho AF, Hawkins BA, Bini LM, Jr PDM & Blackburn TM 2007. Are spatial regression methods a panacea or a Pandora's box? A reply to Beale et al. (2007). *Ecography* 30: 848–851.

Diserud OH & Ødegaard F 2007. A multiple-site similarity measure. *Biology Letters* 3: 20–22.

Dobrovolski R, Melo AS, Cassemiro FAS & Diniz-Filho JAF 2012. Climatic history and dispersal ability explain the relative importance of turnover and nestedness components of beta diversity. *Global Ecology and Biogeography* 21: 191–197.

Dole N, Doledec S & Statzner BS 2012. Spatial autocorrelation patterns of stream invertebrates: exogenous and endogenous factors. *Journal of Biogeography* 39: 56–68.

F Dormann C, M McPherson J, B Araújo M, Bivand R, Bolliger J, Carl G, G Davies R, Hirzel A, Jetz W, Daniel Kissling W, & Kühn I 2007. Methods to account for spatial autocorrelation in the analysis of species distributional data: A review. *Ecography* 30: 609–628.

Dray S, Pelissier R, Couteron P, Fortin M, Legendre P & Peres-Neto P 2012. Community ecology in the age of multivariate multiscale spatial. *Ecological Monographs* 82: 257–275.

Elith J & Leathwick JR 2009. Species distribution models: Ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics* 40: 669–677.

Evans KL, Greenwood JJD & Gaston KJ 2005. Dissecting the species-energy relationship. *Proceedings. Biological sciences/The Royal Society* 272: 2155–2163.

Ewers RM & Didham RK 2008. Pervasive impact of large-scale edge effects on a beetle community. *Proceedings of the National Academy of Sciences of the United States of America* 105: 5426–9.

Ferrier S, Manion G, Elith J & Richardson K 2007. Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. *Diversity and Distributions* 13: 252–264.

Fewster RM & Buckland ST 2001. Similarity indices for spatial ecological data. *Biometrics* 57: 495–501.

Field R, Hawkins B, Cornell H, Currie D, Diniz-Filho A, Guegan JF, Kaufman D, Kerr J, Mittelbach G, Oberdorf T, O'Brien E, & Turner J 2008. Spatial species-richness gradients across scales: a meta-analysis. *Journal of Biogeography* 36: 132–147.

Fischer M & Getis A 2010. Handbook of Applied Spatial Analysis. Springer Berlin Heidelberg.

Fisher JI & Mustard JF 2007. Cross-scalar satellite phenology from ground, Landsat, and MODIS data. *Remote Sensing of the Environment* 109: 261–273.

Fitzpatrick MC, Sanders NJ, Normand S, Svenning JC, Ferrier S, Gove AD & Dunn RR 2013. Environmental and historical imprints on beta diversity: Insights from variation in rates of species turnover along gradients. *Proceedings. Biological sciences/The Royal Society* 280: 20131201.

Fjeldsa J, Lambin E & Ivlertens B 1999. Correlation between endemism and local ecoclimatic stability documented by comparing Andean bird distributions and remotely sensed land surface data. *Ecography* 22: 63–78.

Fjeldså J & Rahbek C 2006. Diversification of tanagers, a species rich bird group, from lowlands to montane regions of South America. *Integrative and comparative biology* 46: 72–81.

Foody GM 2004. Spatial nonstationarity and scale-dependency in the relationship between species richness and environmental determinants for the sub-Saharan endemic avifauna. *Global Ecology and Biogeography* 13: 315–320.

Fortin M & Dale MRT. 2005. Spatial analysis: A guide for ecologists First ed. Cambridge University Press, Cape Town.

Fortin M, James PMA, Mackenzie A, Melles SJ & Rayfield B 2012. Spatial statistics, spatial regression, and graph theory in ecology. *Spatial Statistics* 1: 100–109.

Fotheringham AS, Brundson C & Charlton M 2002. Geographically Weighted Regression. John Wiley & Sons. New York, New York

Francis J & Klopatek J 2000. Multiscale effects of grain size on landscape pattern analysis. *Geographic Information Sciences: A Journal of Association of Chines Professionals in Geographic Information Systems* 6.

Gaston KJ 2000. Global patterns in biodiversity. *Nature* 405: 220–227.

Gaston KJ, Davies RG, Orme CD, Olson VA, Thomas GH, Ding TS, Rasmussen PC, Lennon JJ, Bennett PM, Owens IP & Blackburn TM. 2007. Spatial turnover in the global avifauna. Proceedings. Biological sciences/The Royal Society 274: 1567–1574.

Gentleman R, Hornik K & Parmigiani G 2008. Numerical Ecology with R. Springer.

Getis A 1999. Spatial statistics. In J. W. and S. Longley, M. Goodchild, D. Maguire, and D. Rhind, editor. GIS: Principles, Techniques, Management, and Applications. pp 239–252

Gibson C & Ostrom E 1998. Scaling issues in social sciences. *A report for International Human Dimension Programme on Global Environmental Climate Change*.

Gitelson A, Peng Y, Masek JG, Rundquist DC, Verma S, Suyker A, Baker JM, Hatfield JL & Meyers T 2012. Remote estimation of crop gross primary production with Landsat data. *Remote Sensing of Environment* 121: 404–414.

Goodchild ME & Quattrochi DA. 1997. Scale, multiscaling, remote sensing, and GIS 406: 1–11.

Gosz JR 1992. Gradient analysis of ecological change in time and space: Implications for Forest Management. *Ecological Applications* 2: 248–261.

Gotelli NJ & Colwell RK 2011. Estimating species richness. *Biological diversities: Frontiers in measurement and assessment* 12: 39–54

Graham MH 2003. Confronting multicollinearity in ecological multiple linear regression. *Ecology* 84: 2809–2815.

Griffith DA 2012. Spatial statistics: A quantitative geographer's perspective. *Spatial Statistics* 1: 3–15.

Guerin GR, Biffin E & Lowe A 2013. Spatial modelling of species turnover identifies climate ecotones, climate change tipping points and vulnerable taxonomic groups. *Ecography* 36: 1086–1096.

Guèze M, Paneque-Gálvez C, Luz AC, Pino J, Orta-Martínez M, Reyes-García V & Manuel J 2013. Determinants of tree species turnover in a southern Amazonian rain forest. *Journal of Vegetation Science* 24: 284–295

Guisan A & Thuiller W 2005. Predicting species distribution: Offering more than simple habitat models. *Ecology letters* 8: 993–1009.

Hannan E & Quin B 1979. The determination of the order of an autoregression. *Journal of the Royal Statistical Society* 41: 190–195.

Hawkins BA, Field R, Cornell HV, Currie DJ, Guégan JF, Kaufman DM, Kerr JT, Mittelbach GG, Oberdorff T, O'Brien EM & Porter EE 2003. Energy, water, and broad-scale geographic patterns of species richness. *Ecology* 84: 3105–3117.

Hijmans RJ, Cameron SE, Parra JL, Jones PG & Jarvis A 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25: 1965–1978.

Holt RD, Lawton JH, Polis GA & Martinez ND 1999. Trophic rank and the species-area relationship. *Ecology* 80: 1495–1504.

Houlahan JE, Findlay CS, Schmidt BR, Meyer AH & Kuzmin SL 2000. Quantitative evidence for global amphibian population declines. *Nature* 404: 752–755.

Hui C 2009. On the scaling patterns of species spatial distribution and association. *Journal of theoretical biology* 261: 481–7.

Hui C 2015. Unlocking patterns of nature – the marriage of mathematics and ecology.

Hui C & McGeoch MA. 2014. Zeta diversity as a concept and metric that unifies incidence-based biodiversity patterns. *The American Naturalist* 184:6 84–694.

IUCN 2014. The Festschrift on the 50th threatened Species Anniversary of The IUCN Red List of threatened Species.

Jetz W, Kreft H, Ceballos G & Mutke J 2009. Global associations between terrestrial producer and vertebrate consumer diversity. *Proceedings of the Royal Society B* 276: 269–278.

Jetz W, Rahbek C & Lichstein JW 2005. Local and global approaches to spatial data analysis in ecology. *Global Ecology and Biogeography* 14: 97–98

Jost L 2007. Partitioning diversity into independent alpha and beta components. *Ecology* 88: 2427–2439.

Jurasinski G, Retzer V & Beierkuhnlein C 2009. Inventory, differentiation, and proportional diversity: A consistent terminology for quantifying species diversity. *Oecologia* 159: 15–26.

Jurasinski G, Jentsch A, Retzer V & Beierkuhnlein C 2012. Detecting spatial patterns in species composition with multiple plot similarity coefficients and singularity measures. *Ecography* 35: 73–88.

Kallimanis AS & Koutsias N 2013. Geographical patterns of Corine land cover diversity across Europe: The effect of grain size and thematic resolution. *Progress in Physical Geography* 37: 161–177.

Katayama N, Amano T, Naoe S, Yamakita T, Komatsu I, Takagawa S, Sato N, Ueta M & Miyashita T 2014. Landscape heterogeneity-biodiversity relationship: Effect of range size. *PloS One* 9: e93359.

Keil P, Schweiger O, Kühn I, Kunin WE, Kuussaari M, Settele J, Henle K, Brotons L, Pe'er G, Lengyel S & Moustakas A 2012. Patterns of beta diversity in Europe: The role of climate, land cover and distance across scales. *Journal of Biogeography* 39: 1473–1486.

Kiesecker JM, Blaustein AR & Belden LK 2001. Complex causes of amphibian population declines. *Nature* 410: 681–684.

Kissling W & Carl G 2008. Spatial autocorrelation and the selection of simultaneous autoregressive models. *Global Ecology and Biogeography* 17: 59–71.

Koleff P, Gaston KJ & Lennon JJ 2003. Measuring beta diversity for presence – absence data. *Journal of Animal Ecology* 72: 367–382.

Koleff P, Lennon J & Gaston KJ 2003. Are there latitudinal gradients in species turnover? *Global Ecology and Biogeography* 12: 483–498.

Korhonen JJ 2014. Species turnover of aquatic organisms in space and time Patterns in community composition and diversity.

Krebs CJ 1999. Similarity coefficients and cluster analysis. *Ecological Methodology: Inglaterra*

Lande R 1996. Statistics and partitioning of species diversity, and similarity among multiple communities. *Oikos* 76: 5–13.

Lark RM 2011. Spatially nested sampling schemes for spatial variance components: Scope for their optimization. *Computers and Geosciences* 37: 1633–1641.

Lausch A, Pause M, Merbach I, Zacharias S, Doktor D & Volk M 2013. A new multiscale approach for monitoring vegetation using remote sensing-based indicators in laboratory, field, and landscape. *Environmental monitoring and assessment* 185: 1215–1235.

Le Houérou HN. 2009. Bioclimatology and biogeography of Africa. Berlin: Springer.

Legendre P 1993. Spatial autocorrelation: Trouble or new paradigm. *Ecology* 74: 1659–1673.

Legendre P 2008. Studying beta diversity: Ecological variation partitioning by multiple regression and canonical analysis. *Journal of plant ecology* 1: 3–8.

Legendre P 2014. Interpreting the replacement and richness difference components of beta diversity. *Global Ecology and Biogeography* 23: 1324–1334.

Legendre P, Borcard D, Couteron P, Peres-Neto PR, Tuomisto H, Ruokolainen K, Pelissier R, Dray S & Laliberté E 2005. Analyzing beta diversity: partitioning the spatial variation of community composition data. *Ecological Monographs* 75: 435–450.

Legendre P & Legendre L 1998. Numerical Ecology. Amsterdam. Elsevier ed. Eng.

Legendre P, Mi X, Ren H, Ma K, Yu M, Sun IF & He F 2009. Partitioning beta diversity in a subtropical broad-leaved forest of China. *Ecology* 90: 663–674.

Lennon JJ 2000. Red-shifts and red herrings in geographical ecology. *Ecography*. 101–113.

Lennon JJ, Koleff P, Greenwood JJD & Gaston KJ 2001. The geographical structure of British bird distributions: diversity, spatial turnover and scale. *Journal of Animal Ecology* 70: 966–979.

Levin S 1992. The problem of pattern and scale in ecology. *Ecological Society of America* 73: 1943–1967.

Lichstein J, Simons T, Shriner S & Franzreb K 2002. Spatial autocorrelation and autoregressive models in ecology. *Ecological Monographs* 72: 445–463.

Linder HP, de Klerk HM, Born J, Burgess ND, Fjeldså J & Rahbek C 2012. The partitioning of Africa: Statistically defined biogeographical regions in sub-Saharan Africa. *Journal of Biogeography* 39: 1189–1205.

Loreau M 2000. Are communities saturated? On the relationship between alpha, beta and gamma diversity. *Ecology Letters* 3: 73–76.

Louda SM 1999. Comparisons of related rare and common species in the study of plant rarity. *Conservation Biology* 13: 493–498.

Lustig A, Stouffer DB, Roigé M & Worner SP 2015. Towards more predictable and consistent landscape metrics across spatial scales. *Ecological Indicators* 57: 11–21.

Mac Nally R Mac, Fleishman E, Bulluck LP & Betrus CJ 2004. Comparative influence of spatial scale on beta diversity within regional assemblages of birds and butterflies. *Journal of Biogeography* 31: 917–929.

Mac R, Fleishman E, Bulluck LP, Betrus CJ, Journal S, Jun N, Nally R Mac, Fleishman E, Bulluckt LP & Betrus CJ 2004. Comparative influence of spatial scale on beta diversity within regional assemblages of birds and butterflies. *Journal of Biogeography* 31: 917–929.

MacArthur R 1965. Patterns of species diversity. *Biological reviews* 40: 510–533.

Magurran AE 2004. Introduction: Measurement of biological diversity. Pages 19–70 Measuring Biological diversity. Carlton: Blackwell Publishing.

Magurran AE & Henderson P 2003. Explaining the excess of rare species in natural species abundance distributions. *Letters to nature* 422: 714–716.

Marceau DJ & Hay GJ 1999. Remote sensing contributions to the scale issue. *Canadian Journal of Remote Sensing* 25: 357–366.

Martin B & Ferrer M 2015. Temporally variable environments maintain more beta diversity in Mediterranean landscapes. *Acta Oecologica* 68: 1–10.

Maxime R, Gourlet-fleury S, Couteron P, Nasi R & Thompson JD 2008. Regional variation in tropical forest tree species composition in the Central African Republic: An assessment based on inventories by forest companies. *Journal of Tropical Ecology* 24: 663–674.

McCain CM 2007. Could temperature and water availability drive elevational species richness patterns? A global case study for bats. *Global Ecology and Biogeography* 16: 1–13.

McCain CM & Grytnes JA 2010. Elevational gradients in species richness. *Encyclopaedia of Life Sciences* 12: 1–10.

Mcdonald R, Mcknight M, Weiss D, Selig E, Connor MOÕ, Violin C & Moody A 2005. Species compositional similarity and ecoregions : Do ecoregion boundaries represent zones of high species turnover? *Biological Conservation* 126: 24–40.

McGarigal K 2000. What is a Landscape? *Biological Conservation* 107: 55–69.

McGlynn TP, Weiser MD & Dunn RR 2010. More individuals but fewer species: testing the "more individuals' hypothesis" in a diverse tropical fauna. *Biology letters* 6: 490–493.

McKnight MW, White PS, McDonald RI, Lamoreux JF, Sechrest W, Ridgely RS, Stuart SN 2007. Putting beta-diversity on the map: Broad-scale congruence and coincidence in the extremes. *PLoS Biology* 5: 2424–2432.

Melo AS, Rangel TFLVB & Diniz-filho AF 2009. Environmental drivers of beta-diversity patterns in New-World birds and mammals. *Ecography* 32: 226–236.

Murphy SJ, Audino LD, Whitacre J, Eck JL, Wenzel JW, Queenborough SA & Comita LS 2015. Species associations structured by environment and land-use history promote beta-diversity in a temperate forest. *Ecology* 96: 705–715.

Navarro-sigüenza A & Rodríguez P 2014. Spatial scale and β-diversity of terrestrial vertebrates in Mexico. *Revista mexicana de biodiversidad* 85: 918–930.

Nekola JC & White PS 1999. The distance decay of similarity in biogeography and ecology. *Journal of Biogeography* 26: 867–878.

Nthiwa AN 2011. Modelling scale and effects of the modifiable areal unit problem on multiple deprivations in Istanbul, Turkey. University of Twente.

O'Neill RVO, Turner SJ, Cullinan VI, Coffin DP, Cook T, Conley W & Brunt J 1991. Multiple landscape scales: An intersite comparison. *Landscape Ecology* 5: 137–144.

Ortiz-Yusty C, Páez V & Zapata F 2013. Temperature and precipitation as predictors of species richness in northern Andean amphibians from Colombia. *Caldasia* 35: 65–80.

Oslisly R, White L, Bentaleb I, Favier C, Fontugne M, Gillet J-F & Sebag D 2013. Climatic and cultural changes in the west Congo Basin forests over the past 5000 years. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 368: 201-204.

Parviainen M, Luoto M & Heikkinen R 2010. NDVI-based productivity and heterogeneity as indicators of plant-species richness in boreal landscapes. *Boreal environment research* 6095: 301–318.

Peet RK 1974. The measurement of species diversity. *Annual Review of Ecology and Systematics* 5: 285–307.

Peters DPC, Bestelmeyer BT & Turner MG 2007. Cross – scale interactions and changing pattern – process relationships: Consequences for system dynamics. *Ecosystems* 10: 790–796.

Pettorelli N, Vik JO, Mysterud A, Gaillard J, Tucker CJ, Stenseth NC & Lyon CB 2005. Using the satellite-derived NDVI to assess ecological responses to environmental change. *Trends in ecology & evolution* 20: 503–510.

Phillips LB, Hansen AJ, & Flather CH 2008. Evaluating the species energy relationship with the newest measures of ecosystem energy: NDVI versus MODIS primary production. *Remote Sensing of Environment* 12: 4381–4392

Pineda E & Lobo J 2009. Assessing the accuracy of species distribution models to predict amphibian species richness patterns. *Journal of Animal Ecology* 78: 182–190.

Podani J & Schmera D 2016. Once again on the components of pairwise beta diversity. *Ecological Informatics* 32: 63–68.

Poynton JC 2013. Afrotemperate amphibians in southern and eastern Africa: A critical review. *African Journal of Herpetology* 62: 5–20.

Propastin P, Kappas M & Erasmi S 2008. Application of Geographically Weighted Regression to investigate the impact of scale on prediction uncertainty by modelling relationship between vegetation and climate. *International journal of spatial data infrastructures research* 3: 73–94.

Qi Y & Wu J 1996. Effects of changing spatial resolution on the results of landscape pattern analysis using spatial autocorrelation indices. *Landscape Ecology* 11: 39–49.

Qian H 2009. Global comparisons of beta diversity among mammals, birds, reptiles, and amphibians across spatial scales and taxonomic ranks. *Journal of Systematics and Evolution* 47: 509–514.

Qian H 2010. Environment – richness relationships for mammals, birds, reptiles, and amphibians at global and regional scales. *Ecological Research* 25: 629–637.

Qian H & Ricklefs RE 2012. Disentangling the effects of geographic distance and environmental dissimilarity on global patterns of species turnover. *Global Ecology and Biogeography* 21: 341–351.

Qian H & Shimono A 2012. Effects of geographic distance and climatic dissimilarity on species turnover in alpine meadow communities across a broad spatial extent on the Tibetan Plateau. *Plant Ecology* 8: 1–8.

Qian H & Wang X 2015. Global relationships between beta diversity and latitude after accounting for regional diversity. *Ecological Informatics* 25: 10–13.

Qian H, Wang X, Wang S & Li Y 2007. Environmental determinants of amphibian and reptile species richness in China. *Ecography* 30: 471–482.

Rahbek C 2005. The role of spatial scale and the perception of large-scale species-richness patterns. *Ecology Letters* 8:224–239.

Rahbek C 2006. Diversification of tanagers, a species rich bird group, from lowlands to montane regions of South America. *Integrative and comparative biology* 46: 72–81.

Harvey PH, Colwell RK, Silvertown JW & May RM 1996. Null models in ecology. *Annual Review of Ecology and Systematics* 14: 189–211

Réjou-méchain AM & Hardy OJ 2012. Properties of similarity indices under niche-based and dispersal-based processes in communities. *American Naturalist* 177: 589–604.

Richerson P & Lum K 1980. Patterns of plant species diversity in California: Relation to weather and topography. *The American Naturalist* 116: 504–536.

Ricklefs RE 2007. History and diversity: Explorations at the intersection of ecology and evolution. *The American Naturalist* 17: 56–70.

Ricotta C & Pavoine S 2015. Short communication a multiple-site dissimilarity measure for species presence/absence data and its relationship with nestedness and turnover. *Ecological Indicators* 54: 203–206.

Rodríguez MÁ, Alfonso J & Hawkins BA 2005. Energy, water and large-scale patterns of reptile and amphibian species richness in Europe. *Acta Oecologica* 28: 65–70.

Ronce O 2007. How does it feel to be like a rolling stone? Ten questions about dispersal evolution. *The Annual Review of Ecology, Evolution, and Systematics* 38: 231–253.

Rouquette JR, Dallimer M, Armsworth PR, Gaston KJ, Maltby L & Warren PH 2013. Species turnover and geographic distance in an urban river network. *Diversity and Distributions* 19: 1429–1439.

Roy D 1997. Communication signals and sexual selection in amphibians. *Current Science* 72: 923–927.

Saas Y & Gosselin F 2014. Comparison of regression methods for spatially autocorrelated count data on regularly and irregularly spaced locations. *Ecography* 37: 476–489.

Saito VS, Soininen J, Fonseca-Gessner AA & Siqueira, T 2015. Dispersal traits drive the phylogenetic distance decay of similarity in Neotropical stream metacommunities. J*ournal of Biogeography* 42: 2101–2111

Sale PF 1989. Appropriate spatial scales for studies of reef-fish ecology. J*ournal of Biogeography* 42:2101–2111.

Schall JJ & Pianka ER 1977. Species densities of reptiles and amphibians on the Iberian Peninsula. *Doñana, Acta Vertebrata* 4: 27–34.

Schall JJ & Pianka ER 1978. Geographical trends in number of species. *Science* 201: 679–686.

Schoener TW 2015. Rate of species turnover decreases from lower to higher organisms: A review of the data. *Oikos* 41: 372–377.

Seymour CL, De Klerk HM, Channing AA & Crowe TM 2001. The biogeography of the Anura of sub-equatorial Africa and the prioritisation of areas for their conservation. *Biodiversity and Conservation* 10: 2045–2076.

Shaffer HB, Fellers GM, Magee A &Voss SR 2000. The genetics of amphibian declines: population substructure and molecular differentiation in the Yosemite Toad, Bufo canorus (Anura, Bufonidae) based on single-strand conformation polymorphism analysis (SSCP) and mitochondrial DNA sequence data. *Molecular Ecology* 9: 245–257.

Sheehan KR, Strager MP & Welsh SA 2013. Advantages of Geographically Weighted Regression for modelling benthic substrate in two greater Yellowstone ecosystem streams *Environmental Modelling & Assessment* 18: 209–219.

Shmida A 1984. Measuring beta diversity with presence-absence data. *Journal of Ecology* 72: 1055–1064.

Socolar J, Gilroy J, Edwards DP & Kunin WE 2015. How should beta-diversity inform biodiversity conservation? *Trends in Ecology & Evolution* 31: 67–80.

Sodhi NS, Bickford D, Diesmos AC, Lee TM, Koh LP, Brook BW, Sekercioglu CH & Bradshaw CJ 2008. Measuring the meltdown: Drivers of global amphibian extinction and decline. *PLoS ONE* 3: e1636.

Soininen J, Lennon JJ & Hillebrand H 2007. A multivariate analysis of beta diversity across organisms and environments. *Ecology* 88: 2830–2838.

Stuart SN 2004. Status and trends of amphibian declines and extinctions worldwide. *Science* 306: 1783–1786.

Sykes AO 2007. An introduction to regression analysis. *American Statistician* 4*:* 101–195

Tang Z, Fang J, Chi X, Yang Y, Ma W, Mohhamot A, Guo Z, Liu Y & Gaston KJ 2012. Geography, environment, and spatial turnover of species in China's grasslands. *Ecography* 35: 1103–1109.

Thompson S, Anderson K & Bartlein M 1999. Atlas of Relations Between Climatic Parameters and Distributions of Important Trees and Shrubs in North America. *U.S. Geological Survey Professional Paper* 1650 A&B.

Tokeshi M 1993. Species abundance patterns and community structure. *Advances in ecological research* 24: 111–186.

Tsai PJ 2011. Spatial autocorrelation calculations of the nine malignant neoplasms in Taiwan in 2005-2009: a gender comparison study. *Chinese Journal of Cancer* 30: 757–765.

Tumoisto H & Ruokolainen K 2008. Analyzing or explaining beta diversity? *Ecological Society of Am*erica 89: 3244–3256.

Tuomisto H 2010a. A diversity of beta diversities: Straightening up a concept gone awry. Part 1. Defining beta diversity as a function of alpha and gamma diversity. *Ecography* 33: 2–22.

Tuomisto H. 2010b. A diversity of beta diversities: straightening up a concept gone awry. Part 2. Quantifying beta diversity and related phenomena. *Ecography* 33: 23–45.

Turner MG 1989. Landscape ecology: the effect of pattern on process. *Annual Review of Ecology and Systematics* 20: 171–197.

Ulrich W & Almeida-neto M 2012. On the meanings of nestedness: Back to the basics. *Ecography* 35: 1–7.

Valdujo PH, Carnaval AC & Graham CH 2013. Environmental correlates of anuran beta diversity in the Brazilian Cerrado. *Ecography* 36: 708–717.

Van Rensburg BJ, Koleff P, Gaston KJ & Chown SL 2004. Spatial congruence of ecological transition at the regional scale in South Africa. *Journal of Biogeography* 31: 843–854.

Vasudevan K, Kumar A & Chellam R 2006. Species turnover: The case of stream amphibians of rainforests in the Western Ghats, southern India. *Biodiversity & Conservation* 11: 3515–3525.

Veech JA & Crist TO 2007. Habitat and climate heterogeneity maintain beta diversity of birds among landscapes within ecoregions. *Global Ecology and Biogeography* 16: 650–656.

Vellend M 2001. Do commonly used indices of beta diversity measure species turnover? *Journal of Vegetation Science* 12: 545–552.

Volkov I, Banavar JR, Hubbell SP & Maritan A 2003. Neutral theory and relative species abundance in ecology. *Nature* 424: 1035–1037.

Wang IJ 2013. Examining the full effects of landscape heterogeneity on spatial genetic variation: a multiple matrix regression approach for quantifying geographic and ecological isolation. *Evolution; International Journal of Organic Evolution* 67: 3403–11.

Wang Z, Fang J, Tang Z & Shi L 2012. Geographical patterns in the beta diversity of China's woody plants: The influence of space, environment and range size. *Ecography* 35: 1092–1102.

Werner EE, Skelly DK, Relyea RA & Yurewicz KL 2007. Amphibian species richness across environmental gradients. *Oikos* 116: 1697–1712.

Wheeler D & Tiefelsdorf MT 2005. Multicollinearity and correlation among local regression coefficients in Geographically Weighted Regression. *Journal of Geographical Systems* 7: 161–187.

Whittaker R 1960. Vegetation of the Siskiyou Mountains, Oregon and California. *Ecological Monographs* 30: 279–338.

Whittaker R 1972. Evolution and measurement of species diversity. *International Association for Plant Taxonomy* 21: 213–251.

Whittaker RJ, Willis KJ & Field R 2001. Scale and species richness: Towards a general, theory of species diversity hierarchical. *Diversity* 28: 453–470.

Whitton FJS, Purvis A, Orme CDL & Olalla-Tárraga MÁ 2012. Understanding global patterns in amphibian geographic range size: Does Rapoport rule? *Global Ecology and Biogeography* 21: 179–190.

Wiens JA 1989. Spatial scaling in ecology. *Functional ecology* 3: 385–397.

Wiens JJ & Donoghue MJ 2004. Historical biogeography, ecology and species richness. *Trends in Ecology and Evolution* 19: 639–644.

Williams PH 1996. Mapping variations in the strength and breadth of biogeographic transition zones using species turnover. *Proceedings of the Royal Society B: Biological Sciences* 263: 579–588.

Williams PH, De Klerk HM & Crowe TM 1999. Interpreting biogeographical boundaries among Afrotropical birds: Spatial patterns in richness gradients and species replacement. *Journal of Biogeography* 26: 459–474.

Wilson JB & Meurk CD 2011. The control of community composition by distance, environment and history: A regional-scale study of the mountain grasslands of southern New Zealand. *Journal of Biogeography* 38: 2384–2396.

Winship C 1984. Regression models with ordinal variables. *American sociological review* 49: 512–525.

Woodcock CE & Strahler AH 1987. The factor of scale in remote sensing. *Remote Sensing of Environment* 21: 311–332.

Wu J 2004. Effects of changing scale on landscape pattern analysis: Scaling relations. *Landscape Ecology* 19: 125–138.

Wu J & Li H 2006. Concepts of scale and scaling. Scaling and uncertainty analysis in Ecology: Methods and applications, Springer Netherlands.

Wu J & Qi Y 2000. Dealing with scale in landscape analysis: an overview. *Annals of GIS* 6: 1–5.

WWF 2000. Terrestial ecoregions. [Available online] *http://wwf.panda.org/about_our_earth/ecoregions/ecoregion_list/*

Yuan F & Shouraseni R 2007. Analysis of the relationship between NDVI and climate variables in Minnesota using geographically weighted regression and spatial interpolation. *Proc. of ASPRS 2007 Annual conference*, Tampa, Florida. May. 2007.

Zhang Q, Hou X, Li FY, Niu J, Zhou Y & Ding Y 2014. Alpha, beta and gamma diversity differ in response to precipitation in the inner Mongolia grassland. *PLoS ONE* 9: e93518.

Zietz J, Zietz EN & Sirmans GS 2008. Determinants of house prices: A quantile regression approach. The journal of real estate. *Finance and Economics* 37: 317–333.

Zuo W, Moses ME, West GB, Hou C & Brown JH 2012. A general model for effects of temperature on ectotherm ontogenetic growth and development. *Proceedings of the Royal Society B* 279: 1840–6.

APPENDICES

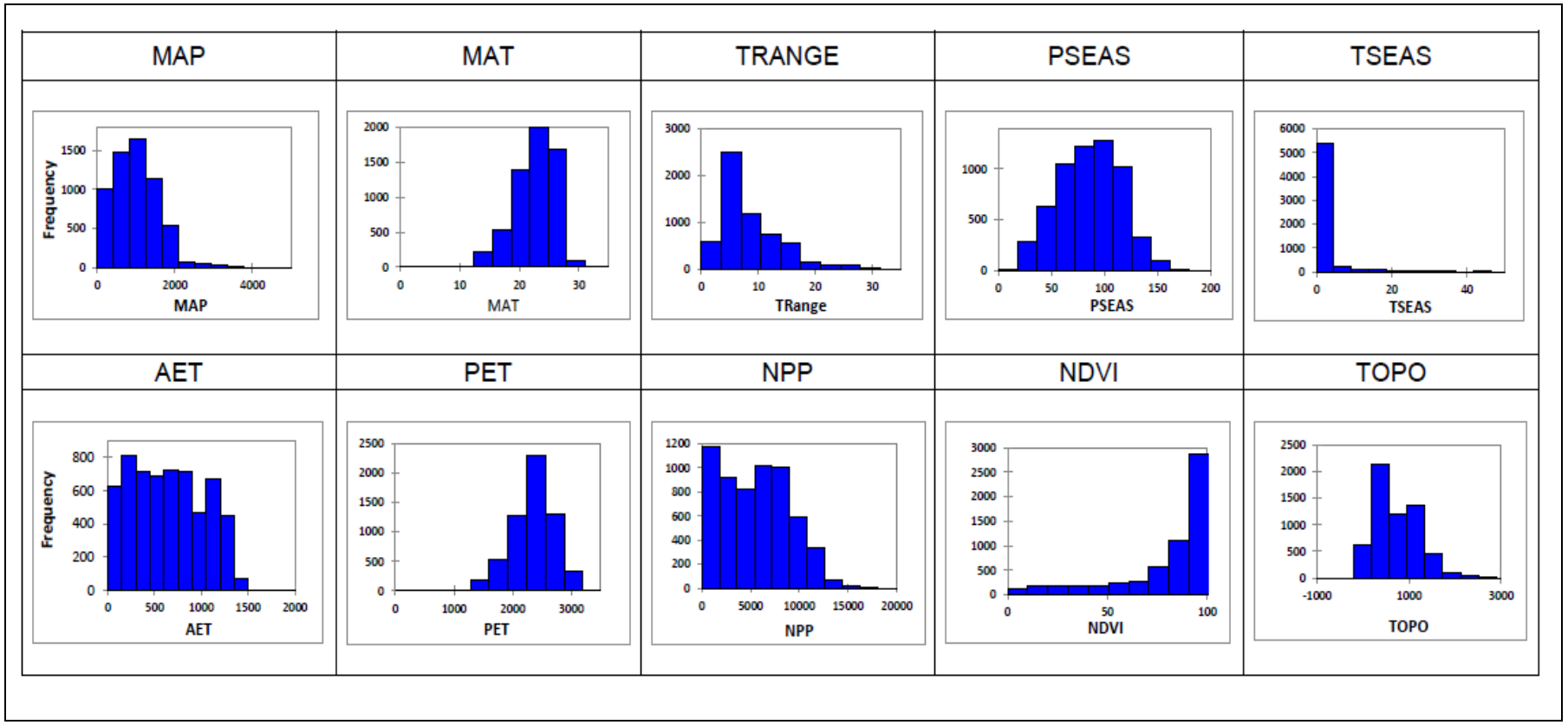
APPENDIX A

R script used to compute zeta diversity partitioning

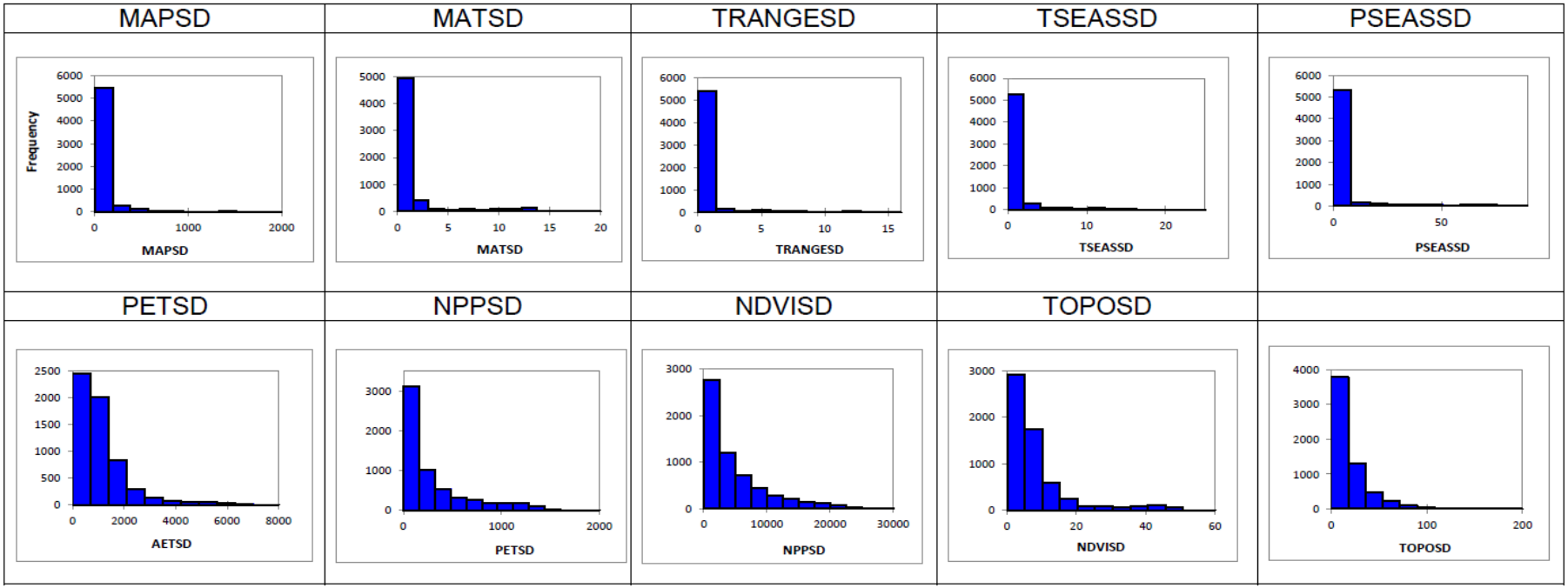
```
rm(list=ls())
graphics.off()
setwd("E:/Students/Users/Postgrad/Curtis/Work/Masters/Masters_Thesis/test/zm/Cang")
install.packages("car")
install.packages("zeta_0.0.0.9000.tar.gz",repos=NULL, type="source")

#loads the zeta diveristy package
require(zeta)

setwd("../")

#this has the neighbourhoods with nrows as grid cells & coloumns as neighbors
cdata <- read.csv("", sep=",",header=T)
head(cdata)
names(cdata)

names(cdata)
head(sdata)

#make zout - values
zout <-data.frame(gridno = sdata[,1],Z1=0, Z2=0, Z3=0.,Z4=0.,Z5=0.,Z6=0.,Z7=0.,Z8=0.,Z9=0.)

#make zsd - sd for zeta
zsd <-data.frame(gridno = sdata[,1],Z1=0, Z2=0, Z3=0.,Z4=0.,Z5=0.,Z6=0.,Z7=0.,Z8=0.,Z9=0.)

#iterate through all records, for selected neighbourhoods defined by pointdistance
#in ArcMap and exported using python script to ZM_out.csv, determine z for
#neighbourhood with j cells

ptm <- proc.time()
for (i in 1:nrow(cdata))
    {
    y<-unlist(cdata[i,])
    z<-y[!is.na(y)]
    species.data<-sdata[z,2:726]
    species.data[is.na(species.data)]<-0
    #j<-7
    #if (j<=length(z))
    for (j in 1:length(z))
        {
        zeta<-Zeta.order(species.data,order = j)
        zout[i,j+1]<-zeta$val
        zsd[i,j+1]<-zeta$sd
        }
    }
proc.time() - ptm

# Write to a file, suppress row names
# instead of "NA", output blank cells
write.csv(zout, "", row.names=FALSE, na="")
write.csv(zsd, "", row.names=FALSE, na="")
#paste0("Z",j)
```

APPENDIX B

Histograms computed for predictor variables



Histograms computed for predictor variables. An asterisk indicates variables that were transformed (*)



Histograms computed for predictor variables. Variables that were transformed are indicated by an asterisk (*)